PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C07K 14/52, A61K 38/19 | A1

(11) International Publication Number: **WO 98/34951**

(43) International Publication Date: 13 August 1998 (13.08.98)

(21) International Application Number: PCT/AU98/00078

(22) International Filing Date: 11 February 1998 (11.02.98)

(30) Priority Data:

PO 5067	11 February 1997 (11.02.97)	AU
PO 6420	24 April 1997 (24.04.97)	AU
PO 8963	3 September 1997 (03.09.97)	AU
PP 0961	16 December 1997 (16.12.97)	AU

(71) Applicant *(for all designated States except US)*: AMRAD OPERATIONS PTY. LTD. [AU/AU]; 576 Swan Street, Richmond, VIC 3121 (AU).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: HILTON, Douglas, J. [AU/AU]; 244 Research Road, Warrandyte, VIC 3113 (AU). STANLEY, Edouard, G. [AU/AU]; 5/150 Barkly Street, North Fitzroy, VIC 3068 (AU). HARVEY, Richard, P. [AU/AU]; 5 George Street, Brunswick, VIC 3050 (AU). BIBEN, Christine [AU/FR]; 6/43 Bank Street, Ascot Vale, VIC 3032 (AU). FABRI, Louis [AU/AU]; 8 Laver Court, Mill Park, VIC 3082 (AU). LAH, Maria [AU/AU]; 6 Kyora Parade, North Balwyn, VIC 3104 (AU). NASH, Andrew, D. [AU/AU]; 24 Green Street, Northcote, VIC 3070 (AU).

(74) Agents: HUGHES, E., John, L. et al.; Davies Collison Cave, 1 Little Collins Street, Melbourne, VIC 3000 (AU).

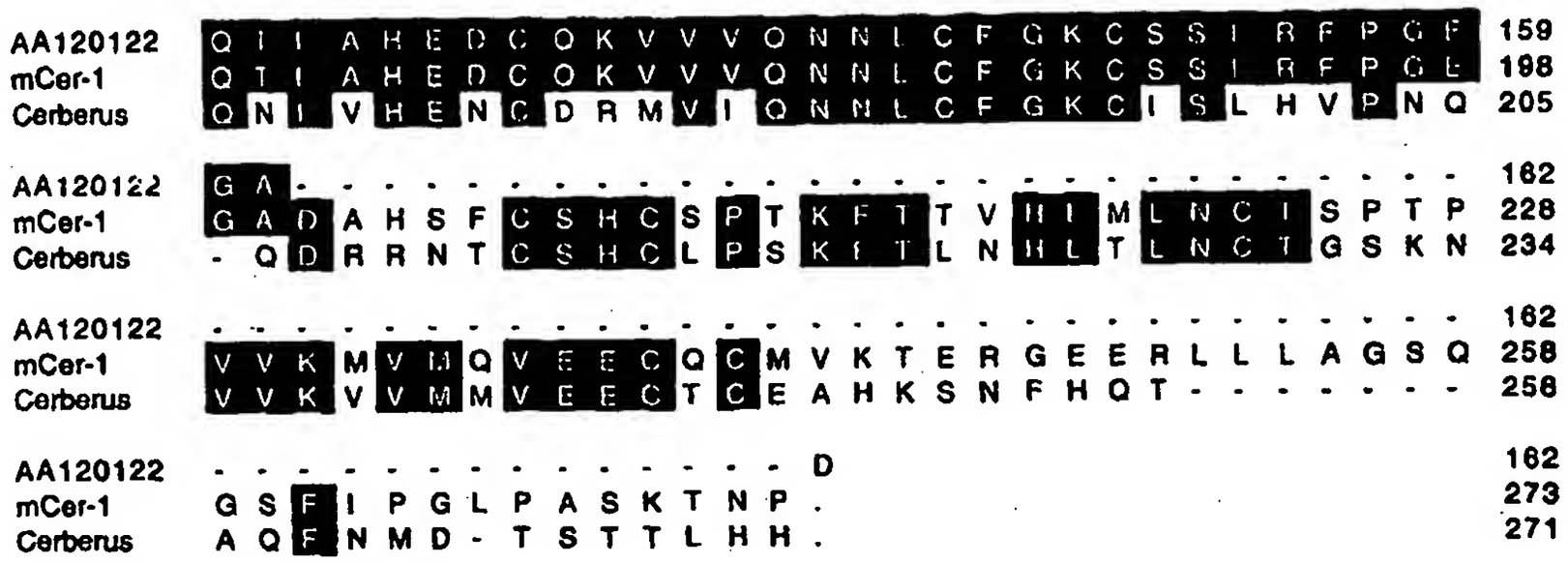
(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: A NEW CYTOKINE FAMILY AND USES THEREOF



(57) Abstract

The present invention relates generally to a new family of cytokine molecules. More particularly, the present invention provides mammalian cytokines which constitute a novel family of cytokines and which are useful in a range of therapeutic and diagnostic applications.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

A NEW CYTOKINE FAMILY AND USES THEREOF

FIELD OF THE INVENTION

The present invention relates generally to a new family of cytokine molecules. More particularly, the present invention provides mammalian cytokines which constitute a novel family of cytokines and which are useful in a range of therapeutic and diagnostic applications.

Bibliographic details of the publications numerically referred to in this specification are collected at the end of the description. Sequence Identity Numbers (SEQ ID NOs.) for the nucleotide and amino acid sequences referred to in the specification are defined following the bibliography.

BACKGROUND OF THE INVENTION

The increasing sophistication of recombinant DNA technology is greatly facilitating research and development in the medical and allied health fields. This is particularly the case in the area of cytokine research.

Cytokines represent an important class of proteinaceous molecules involved in regulation of a vast array of functions in animals including survival, growth, differentiation and effector function of tissue cells. Cytokines generally fall into particular classes encompassing interleukins, colony-stimulating factors, lymphokines, monokines and interferons amongst many others. The identification of new families of cytokines is an important step in facilitating the use of cytokines in therapy and diagnosis.

In accordance with the present invention, a new family of cytokines has been identified comprising mammalian homologues of a protein designated "cerberus" from *Xenopus laevis* embryos (1). The cerberus molecule is a 270 amino acid protein with a role in inducing head structures in the *Xenopus* embryo (1).

SUMMARY OF THE INVENTION

Throughout this specification and the claims which follow, unless the context requires otherwise, the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated integer or group of integers but not the exclusion of any other integer or group of integers.

One aspect of the present invention provides an isolated polypeptide of mammalian origin comprising a signal sequence and a domain conforming to a cystine knot and optionally a long N-terminal domain between said signal sequence and cystine knot domain or a derivative of said polypeptide.

Another aspect of the present invention is directed to an isolated polypeptide of mammalian origin derivative thereof comprising a signal sequence and a domain conforming to the criteria for a cystine knot and optionally a long N-terminal domain between said signal sequence and said cystine knot domain, said polypeptide comprising the amino acid sequence:

$$C\{AA\}Q\{AA\}H\{AA\}C\{AA\}[X^1]_nQN\{AA\}C\{AA\}G\{AA\}C\{AA\}S\{AA\}P$$

wherein

{AA} is an amino acid sequence comprising from about 0 to about 50 amino acid residues;

X^1 is V or I; and

n is 0 or 1.

Yet another aspect of the present invention relates to an isolated polypeptide having the following characteristics:

(i) being glycosylated in its naturally occurring form;

(ii) being secretable in its naturally occurring form;

(iii) comprising a signal sequence and a domain conforming to the criteria for a cystine knot,

said cystine knot domain comprising the sequence:

$$CxTxPFxQxIxHExCxxxVxQNNLCFGKCxSxxxPx_{n1}CSHCxP \text{ [SEQ ID NO:1]}$$

wherein x is any amino acid residue and n_1 is from about 6 to about 10 or comprises a sequence in the cystine knot domain having at least 50% identity to SEQ ID NO:1 excluding the cystine and x residues.

Preferably, the polypeptide further comprises a long N-terminal domain between said signal sequence and said cystine knot.

Still another aspect of the present invention provides an isolated polypeptide having the following characteristics:

(i) being glycosylated in its naturally occurring form;

(ii) being secretable in its naturally occurring form;

(iii) comprising a signal sequence and a domain conforming to the criteria for a cystine knot,

said cystine knot domain comprising the sequence:

CRTVPFNQTIAHEDCQKVVVQNNLCFGKC [SEQ ID NO:2]

or a sequence having at least 45% similarity to SEQ ID NO:2.

Preferably, the polypeptide further comprises a long N-terminal domain between said signal sequence and said cystine knot.

Yet still another aspect of the present invention is directed to a novel polypeptide or a modification to a previously known polypeptide characterised by having a cystine knot domain comprising the sequence:

$$CXXXXXXXXXXHX_aCXXXXXXXXXCXGXCXXX_{n2}CXXCXP$$

wherein:

X is any amino acid residue;

a is from about 2 to about 4; and

$_{n2}$ is from about 6 to about 100.

Even yet another aspect of the present invention contemplates an isolated secretable polypeptide or derivative thereof comprising an amino acid sequence having at least 20% homology to the cerberus protein from *Xenopus laevis* defined in Figure 1.

Another aspect of the present invention is directed to an isolated secretable polypeptide comprising a sequence of amino acids substantially as set forth in SEQ ID NO:4 or having at least 50% similarity thereto and which polypeptide has the identifying characteristics of a CRP cytokine. The molecule defined in SEQ ID NO:4 is from a mouse and is referred to herein as "mCRP-1".

In a related embodiment, there is provided an isolated secretable polypeptide comprising a sequence of amino acids substantially as set forth in SEQ ID NO:5 or having at least 50% similarity thereto and which polypeptide has the identifying characteristics of a CRP cytokine. The molecule defined in SEQ ID NO:5 is from a rat and is referred to herein as "rCRP-2".

Another aspect of the present invention relates to an isolated, secretable polypeptide comprising a sequence of amino acids substantially as set forth in SEQ ID NO:6 or having at least 50% similarity thereto and which polypeptide has the identifying characteristics of a CRP cytokine. The molecule defined in SEQ ID NO:6 is from a mouse and is referred to herein as "mCRP-2".

Yet another aspect of the present invention is directed to a secretable polypeptide comprising a sequence of amino acids substantially as set forth in SEQ ID NO:7 or having at least 50% similarity thereto and which polypeptide has the identifying characteristics of a CRP cytokine. The molecule defined in SEQ ID NO:7 is from a human and is referred to herein as "hCRP-2".

Still yet another aspect of the present invention provides a secretable polypeptide comprising a sequence of amino acids substantially as set forth in SEQ ID NO:19 and/or 20 or having at least 50% similarity thereto and which polypeptide has the identifying characteristics of a CRP cytokine. The molecule defined by SEQ ID NO:19 and/or 20 is from a human and is referred to herein as "hCRP-1".

Even yet another aspect of the present invention relates to a secretable polypeptide comprising a sequence of amino acids substantially as set forth in SEQ ID NO:22 or having at least 50% similarity thereto and which polypeptide has the identifying characteristics of a CRP cytokine. The molecule defined by SEQ ID NO:22 is hCRP-1 but without the intron and corresponding amino acid translation.

Another aspect of the present invention is directed to a nucleic acid molecule comprising a sequence of nucleotides or a complementary sequence of nucleotides which encodes an amino acid sequence selected from SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:19, SEQ ID NO:20 and SEQ ID NO:22 or a sequence having at least 50% similarity thereto.

Yet another aspect of the present invention provides a nucleic acid molecule comprising a nucleotide sequence substantially as set forth in SEQ ID NO:3, SEQ ID NO:18 or SEQ ID NO:27 or a sequence having at least 50% similarity thereto or a sequence capable of hybridizing to one of the above under low stringency conditions at 42°C.

Still yet another aspect of the present invention contemplates a method for modulating expression of a CRP cytokine in a mammal such as a human, primate or laboratory test animal, said method comprising contacting a gene encoding said CRP cytokine with an effective amount of a modulator of CRP cytokine expression for a time and under conditions sufficient to up-regulate or down-regulate or otherwise modulate expression of the CRP cytokine.

Even yet another aspect of the present invention is directed to a method for modulating activity of the CRP cytokine in a mammalian such as a human, said method comprising administering

to said mammal a modulating effective amount of a molecule for a time and under conditions sufficient to increase or decrease CRP cytokine activity. The molecule may be a proteinaceous molecule or a chemical entity and may also be a derivative of a CRP cytokine or its ligand (eg. a receptor) or a chemical analogue or truncation mutant of a CRP cytokine or its ligand.

Still even yet another aspect of the present invention relates to a pharmaceutical composition comprising one or more CRP cytokines or derivatives thereof or a modulator of CRP cytokine expression or CRP cytokine activity and one or more pharmaceutically acceptable carriers and/or diluents. These components are referred to as the active ingredients.

In even a further aspect of the present invention, there is provided a method for detecting CRP in a biological sample from a subject said method comprising contacting said biological sample with an antibody specific for a CRP(or group of CRP s) or its derivatives or homologues for a time and under conditions sufficient for an antibody-CRP complex to form, and then detecting said complex.

Another aspect of the present invention provides a use of a CRP cytokine or its functional derivatives in the manufacture of a medicament for the treatment of defective or deficient CRP mediated activities.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 is a representation of amino acid alignments using Clustal method with PAM250 residue weight table. (a) Alignment of EST AA120122, mCER-1 (mCRP-1) and cerberus (Figure 1); (b) Alignment of mCER-1 (mCRP-1), cerberus and murine DAN (mCRP-2); (c) Alignment between (NDP) Norrie disease protein and murine DAN (mCRP-2); (d) Alignment between mCER-1 (mCRP-1) and murine DAN (mCRP-2); (e) Alignment between human, mouse and rat DAN (hCRP-2, mCRP-2 and rCRP-2, respectively); (f) Alignment between cerberus and mCER-1 (mCRP-1).

Figure 2 is the nucleotide and amino acid sequence of mCRP-1.

Figure 3 is the deglycosylation and Western analysis of secreted mouse mCRP-2 protein. Supernatant from COS cells transfected with a murine mCRP-2-FLAG expression vector was chromatographed over an anti-FLAG affinity column according to manufacturer's instructions (Eastman Kodak Company, New Haven CT). Monomeric CRP protein was isolated by gel filtration and subject to N-linked deglycosylation using N-Glycosidase-F (Boehringer Mannheim, Mannheim, Germany). Purified mCRP-2 protein prior to and after deglycosylation was subject to SDS-PAGE gel and visualised by silver staining and Western blot using an anti-FLAG antibody.

Figure 4 is a graphical representation showing transient DAN expression in cos cells.

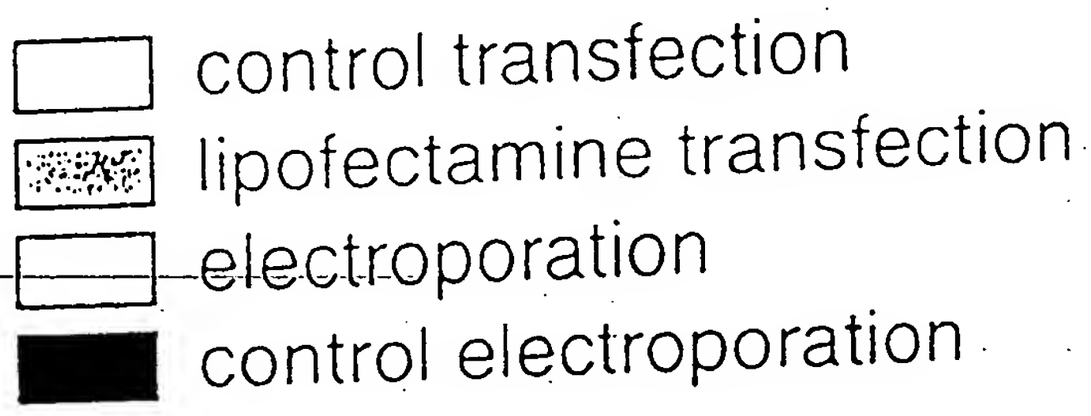


Figure 5 shows the steps in the purification of mCRP-2 (C-FLAG) using affinity chromatography.

Figure 6 is a photographic representation showing SDS-PAGE analysis of mCRP-2 (C-FLAG).

Figure 7 is a graphical representation showing size exclusion analysis of mCRP-2 (C-FLAG).

Figure 8 is a photographic representation of Western Blot analysis of mCRP-1 post M2 affinity purification.

Figure 9 is a representation of SDS PAGE, Western blot and HPLC analysis of purified protein of mCRP-1. (A) SDS PAGE of mCRP-1 following removal of carbohydrate; (B) size exclusion analysis of mCRP-1; (C) regressional analysis of peak fractions containing mCRP-1 compared

to standards; (D) N-terminal amino acid sequence analysis of RP-HPLC purified mCRP-1 [SEQ ID NO:23].

Figure 10 is a representation showing an SDS PAGE, Western blot and HPLC analysis of mCRP-2. (A) SDS-PAGE-Western blot with and without carbohydrate molecules; (B) SDS PAGE- Silver stain either with or without reduction; (D) and (C) size exclusion analysis followed by linear regressional analysis of peak fractions containing mCRP-2; (E) N-terminal amino acid sequencing of mCRP-2 [SEQ ID NO:24].

Figure 11 is a representation showing the complete genomic sequence of hCRP-1 and corresponding amino acid sequences to exons 1 and 2.

Figure 12 is a representation showing the coding sequence of hCRP-1 without the intron and corresponding amino acid translation.

Figure 13 is a representation showing alignment of the predicted amino acid sequence of hCRP-1, mCRP-1 and *Xenopus cerberus*. The alignment was performed using DNA megalign under default conditions. hCRP-1 shares approximately 68% identity with mCRP-1 and approximately 25% identity with *Xenopus cerberus*.

Figure 14 is a representation showing a comparison of mCRP-1 and cerberus activities in *Xenopus* animal cap assays. A. Formation of cement glands in individualised animal caps (stage 35) after mock injection (panel A), or injection of mRNAs encoding mCRP-1 (panel B), CFLAG-mCRP-1 (panel C) or cerberus (panel D). A single darkly pigmented cement gland is induced by injected cerberus mRNAs in some cases (see Figure 3B), but not by mock injection. Note in panel B the secretion of sticky exudate. B. Histogram depicting frequency of cement gland induction in animal caps after injection of mRNAs encoding cerberus, mCRP-1 or CFLAG-mCRP-1, compared to uninjected controls. Data were compiled for each mRNA after examination of 100-120 injected caps. C. RT-PCR analysis of markers induced in animal caps at stage 28 after injection of mRNAs encoding cerberus (XCer; lane 1), mCRP-1 (mCer; lane 2) and BMP4 (lane 5), or after coinjection of 1:1 cerberus: BMP4 (lane 3), and 1:1mCRP-

1:BMP4 (lane 4), compared to uninjected control caps (lane 7) and whole stage 25 embryos (lane 6). *Nkx*2.3 and *Nkx*2.5 are expressed in cardiac progenitors and anterior pharyngeal endoderm. T4 globulin is specific to blood, a ventral mesodermal derivative . *XeHAND* marks cardiac and vascular smooth muscle progenitors in lateral mesoderm. *NCAM* is a pan-neural marker. Otx2 is a marker of anterior tissues, expressed in midbrain, forebrain, placodes, cement gland and anterior mesoderm. Krox20 is expressed in rhombomeres 3 and 5 in the hindbrain. CG13 is expressed specific to cement gland. Edd expression is ubiquitous at low levels but enriched in endoderm. Equal mRNA was assessed by expression of EF-1 alpha (EF1a).

Figure 15 is a photographic representation showing expression of mCRP-1 during gastrulation. A-C, H and I show lateral views of embryos where anterior is to the left, posterior to the right; the arrowhead represents the embryonic/extra-embryonic junction. A. Three early primitive streak stage embryos showing mCRP-1 expression in a midline stripe on the anterior side (left embryo), then progressively (middle and right embryos) in migrating mesendodermal wings arising in the anterior region of the streak. B. Early streak embryo hybridised with probes for both mCRP-1 and *brachyury*. Nascent mesodermal wings positive for *brachyury* are seen on the posterior side, while the mCRP-1 strip is seen anteriorly. At that stage, most or all of the anterior endoderm is of the visceral lineage. C. Late primitive stage embryos hybridised with an mCRP-1 probe. Expression is confined to anterior mesendoderm. D. Distal view of the same embryo as in C, showing lack of mCRP-1 expression in the region of the node. E. Anterior view of the same embryo as in C, showing mCRP-1 expression set back from the embryonic/extra-embryonic junction and absent from the cardiac progenitor region (brackets). F. Early neurula embryo (anterior view) hybridised with a *BMP*2A probe. Expression occurs in the domain of the cardiac progenitors (brackets), but also somewhat into the extra-embryonic domain. G. Neurula embryo (anterior view) hybridised with *Nkx*2-5, specific to the cardiac progenitors at that stage (brackets). H. Sagittal section through a late streak embryo highlighting expression of mCRP-1 in anterior mesendoderm. The arrowhead indicates the position of the node, where expression is absent. I. Enlargement of a section near adjacent to the one shown in H, depicting mCRP-1 expression in both endoderm and associated mesoderm. Note that anterior mesendoderm lacks expression. J. Headfold stage embryo (anterior view)

showing mCRP-1 expression in a wedge of anterior mesendoderm. K. Transverse section through the embryo in J showing mCRP-1 expression in anterior mesendoderm underlying the neural plate, and excluded from the more lateral cardiogenic region (brackets). Abbreviations: en, endoderm; ep, epiblast; me, mesoderm; np, neural plate.

Figure 16 is a photographic representation showing expression in anterior axial mesoderm. Transverse histological sections of an ~E7.75 embryo hybridised in wholemount with an mCRP-1 probe at the level of A. prechordal mesoderm; B. notochordal precursos; C. node. Expression is seen in prechordal and notochordal plates, but is absent from the note. Arrows delimit the apparent extent of these structures. The node is clearly recognised by the recessed nature of its innermost cells (the "pit"). Notochordal and prechordal plate cells have a morphology distinct from surrounding endoderm, consistent with their having a smaller surface area ventrally when viewed by scanning electron microscopy. Prechordal plate is wider than notochordal precursors, forming a wedge shape. Note mCRP-1 expression is head mesoderm.

Figure 17 is a photographic representation showing expression of mCRP-1 in paraxial mesoderm. A. Early headfold stage embryo showing the mCRP-1 anterior expression domain (left; see Figure 15), now faded almost completely, and the first appearance of two strips within paraxial mesoderm. B. Ventroanterior view of a late headfold stage embryo showing absence of mCRP-1 expression in anterior mesendoderm, but stronger expression in paraxial mesoderm. C. Dorsal view of the tail region of a E9.5 embryo (anterior to the left) showing mCRP-1 expression in four stripes in paraxial mesoderm. D. Transverse section through an E8.5 embryo showing that most or all cells with a paraxial strip express mCRP-1. E. Paraxial section through the tail of an E9.5 embryo (anterior to the left) showing all four mCRP-1 strips in paraxial mesoderm. The weaker anterior - most strips mark the rostral region of the most recently formed somites, while the stronger posterior strips are within the presomitic mesoderm. Definitive somite boundaries are indicated by solid arrows. The open arrow indicates the poorly condensed boundary between the forming somite and adjacent presomitic mesoderm. F. E12.5 embryo showing mCRP-1 expression only within nascent and newly formed somites within the tail. Abbreviations: nt: neutral tube; s:somites.

Figure 18 is a photographic representation showing expression in *Otx-2*$^{-/-}$ embryos. The figures shows three *Otx-2*$^{-/-}$ embryos harvested at E6.7 (embryo on the left) and E7.5. In the two most affected (and younger) embryos at the left and centre of the panel (see Ang *et al* 1996), mCER-1 expression (arrows) is seen on the anterior side, but only distally, while in the less affected embryo (right), it extends more toward the embryonic/extra-embryonic junction, as in normal embryos (see Figure 4). Abbreviations: A, anterior; em, embryonic region; ex, extra-embryonic region; P, posterior.

Figure 19 is a representation showing that mCRP-1 maps to the central region of mouse chromosome 4 as determined by interspecific backcross analysis. A. Segregation patterns of mCRP-1 and linked genes in the 92 backcross animals that were typed for all loci. Each column represents the chromosome identified in backcross progeny that was inherited from the (C57BL/6Jx*M. spretus)* F1 parent. The shaded boxes represent the presence of the C57BL/6J allele and the white boxes represent the presence of the *M. spretus* allele. The number of offspring inheriting each type of chromosome is listed at the bottom of each column. B. Partial chromosome 4 linkage map showing the location of mCRP-1 in relation to linked genes. Recombination distances between loci in cM are shown to the left of the chromosome and the positions of loci in human chromosomes, where known, are shown on the right. References for the human map positions of loci cited in this study can be obtained from GDB (Genome Data Base), a computerised database of human linkage information maintained by the William H. Welch Medical Library of The John Hopkins University (Baltimore, MD). C. Southern blot analysis of genomic DNAs from wild type (*Wt),* and those carrying the mutant alleles headblebs (heb), meander tail (mea), polysyndactyly (ps) and pintail (pt). These mutants map to chromosome 4 in the vicinity of mCRP-1. Restriction endonuclease digestion was with *Eco*RI or *Bam*HI. A schematic representation of the mCRP-1 locus is indicated below the panel. The map was determined by restriction enzyme and sequence analysis of clones isolated from a genomic library of the 129 strain. Restriction enzyme sites relevant to this study are indicated. Coding exons of the mCRP-1 gene are boxed. The star indicates the *Eco*RI site absent from the mCRP-1 allele of pt.

Figure 20 is a representation showing: A. Western blot analysis of CFLAG-mCRP-1 protein

purified from CHO cells (see Materials and Methods) before (-) and after (+) treatment with N-glycosidase F (n-gly). B. Western blot analysis of CFLAG-mCRP-1 protein secreted from 293T cells with (+) and without (-) reduction with 100mM dithiothreitol (DTT). mCRP-1 protein was detected with M2 anti-FLAG antibody. The mobility of molecular weight standards (size shown in kilodaltons) is indicated on the left of each panel.

Figure 21 is a photograph representation of a Western blot analysis of 293T hCRP-1-I-SPY transient expression using I-SPY antibody (D11).

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

One aspect of the present invention provides an isolated polypeptide of mammalian origin comprising a signal sequence and a domain conforming to the criteria for a cystine knot and optionally a long N-terminal domain between said signal sequence and cystine knot domain or a derivative of said polypeptide.

The polypeptide of the present invention is, in its naturally occurring form, secretable and glycosylated. The present invention extends, however, to recombinant, synthetic or other modified forms which have an altered glycosylation pattern and/or which have an altered capacity to be secreted from a cell.. For example, the signal sequence may be removed or otherwise inactivated or a signal sequence from another molecule may be fused to the subject polypeptide. The term "signal sequence" is used in its broadest sense and includes a hydrophobic leader sequence. The term "cystine knot" is conveniently as described by McDonald and Hendrickson (7) and Isaacs (8).

Reference herein to a "long" N-terminal domain means that the domain is longer relative to comparable molecules. For example, for convenience, length may be compared relative to Norrie disease protein (NDP) or *D*ifferential screening-selected gene *A*berrative in *N*euroblastoma (DAN[2]). Preferably, the long N-terminal domain when present in the polypeptide of the present invention is from about 100 to about 300 amino acids in length and more preferably from about 100 to about 200 amino acids in length.

In a particularly preferred embodiment of the present invention, the subject polypeptide comprises the long N-terminal domain flanked by the signal sequence and cystine knot domain.

Another aspect of the present invention provides an isolated polypeptide of mammalian origin derivative thereof comprising a signal sequence and a domain conforming to the criteria for a cystine knot and optionally a long N-terminal domain between said signal sequence and said cystine knot domain, said polypeptide comprising the amino acid sequence:

C{AA}Q{AA}H{AA}C{AA}[x^1]$_n$QN{AA}C{AA}G{AA}C{AA}S{AA}P

wherein

{AA} is an amino acid sequence comprising from about 0 to about 50 amino acid residues;

X^1 is V or I; and

n is 0 or 1.

Yet another aspect of the present invention is directed to an isolated polypeptide having the following characteristics:

(i) being glycosylated in its naturally occurring form;

(ii) being secretable in its naturally occurring form;

(iii) comprising a signal sequence and a domain conforming to the criteria for a cystine knot, said cystine knot domain comprising the sequence:

CxTxPFxQxIxHExCxxxVxQNNLCFGKCxSxxxPxn_1CSHCxP [SEQ ID NO:1]

wherein x is any amino acid residue and n_1 is from about 6 to about 10 or comprises a sequence in the cystine knot domain having at least 50% identity to SEQ ID NO:1 excluding the cystine and x residues.

Preferably, the isolated polypeptide comprises an N-terminal domain of from about 100 to about 200 amino acids between the signal sequence and the cystine knot domain.

Preferably, the percentage identity of related molecules to SEQ ID NO:1 is at least about 55%, more preferably at least about 60%, still more preferably at least about 65%, even more preferably at least about 70% or greater such as at least about 71-75%, 76-80%, 81-85%, 86-90% or 91-100%.

According to a particularly preferred embodiment of the present invention, there is provided an isolated polypeptide having the following characteristics:

(i) being glycosylated in its naturally occurring form;

(ii) being secretable in its naturally occurring form;

(iii) comprising a signal sequence and a domain conforming to the criteria for a cystine knot, said cystine knot domain comprising the sequence:

CRTVPFNQTIAHEDCQKVVVQNNLCFGKC [SEQ ID NO:2]

or a sequence having at least 45% similarity to SEQ ID NO:2.

Preferably, the isolated polypeptide comprises a N-terminal domain of from about 100 to about 200 amino acids between the signal sequence and the cystine knot.

Preferably, the percentage identity of related molecules to SEQ ID NO:2 is at least about 55%, more preferably at least about 60%, still more preferably at least about 65%, even more preferably at least about 70% or greater such as at least about 71-75%, 76-80%, 81-85%, 86-90% or 91-100%.

The present invention further extends to novel polypeptides or modifications to previously known polypeptides characterised by having a cystine knot domain comprising the sequence:

$$CXXXXXXXXXXXHX_aCXXXXXXXXXXCXGXCXXX_{n2}CXXCXP$$

wherein:

X is any amino acid residue;

a is from about 2 to about 4; and

$_{n2}$ is from about 6 to about 100.

Particularly preferred embodiments of these aspects of the present invention comprise a long N-terminal domain flanked by a signal sequence and a sequence conforming to the criteria for a cystine knot.

Preferably, according to this aspect of the present invention, the sequences of part of the cystine knot domain is:

$$CXXXXXTQXXXHXXCXXX'XIQNXXCXGXCXSXXVPNX_{n3}CXXCXP$$

wherein:

X is any amino acid residue;

X′ is preferably K; and

$_{n3}$ is from about 6 to about 13.

Another aspect of the present invention provides an isolated secretable polypeptide or derivative thereof comprising an amino acid sequence having at least 20% similarity to the cerberus protein from *Xenopus laevis* defined in Figure 1.

It is proposed, in accordance with the present invention, that the secretable polypeptide as hereinbefore defined above constitutes a novel family of cytokines. The present invention extends, therefore, to all members as defined to be part of this family and to derivatives thereof.

For convenience, the novel family of cytokines is hereinafter referred to as the cerberus related protein (CRP) family of cytokines. Individual members are defined by a lower case letter prefix to CRP to indicate the mammalian origin. Where more than one CRP cytokine has been identified from a particular species, the abbreviation is followed by a number. For example

"rCRP" refers to a CRP cytokine from a rat, "hCRP" refers to a human CRP and "mCRP-1" and "mCRP-2" are murine CRPs.

In a particular embodiment, the present invention further provides an isolated secretable polypeptide comprising a sequence of amino acids substantially as set forth in SEQ ID NO:4 or having at least 50% similarity thereto and which polypeptide has the identifying characteristics of a CRP cytokine. The molecule defined in SEQ ID NO:4 is from a mouse and is referred to herein as "mCRP-1".

In a related embodiment, the present invention provides an isolated secretable polypeptide comprising a sequence of amino acids substantially as set forth in SEQ ID NO:5 or having at least 50% similarity thereto and which polypeptide has the identifying characteristics of a CRP cytokine. The molecule defined in SEQ ID NO:5 is from a rat and is referred to herein as "rCRP-2".

A further embodiment of the subject invention provides an isolated, secretable polypeptide comprising a sequence of amino acids substantially as set forth in SEQ ID NO:6 or having at least 50% similarity thereto and which polypeptide has the identifying characteristics of a CRP cytokine. The molecule defined in SEQ ID NO:6 is from a mouse and is referred to herein as "mCRP-2".

Still a further embodiment provides a secretable polypeptide comprising a sequence of amino acids substantially as set forth in SEQ ID NO:7 or having at least 50% similarity thereto and which polypeptide has the identifying characteristics of a CRP cytokine. The molecule defined in SEQ ID NO:7 is from a human and is referred to herein as "hCRP-2".

Even yet another embodiment of the present invention is directed to a secretable polypeptide comprising a sequence of amino acids substantially as set forth in SEQ ID NO:19 and/or 20 or having at least 50% similarity thereto and which polypeptide has the identifying characteristics of a CRP cytokine. The molecule defined by SEQ ID NO:19 and/or 20 is from a human and is referred to herein as "hCRP-1".

Another embodiment of the present invention provides a secretable polypeptide comprising a sequence of amino acids substantially as set forth in SEQ ID NO:22 or having at least 50% similarity thereto and which polypeptide has the identifying characteristics of a CRP cytokine. The molecule defined by SEQ ID NO:22 is hCRP-1 but without the intron and corresponding amino acid translation.

Preferably, with respect to the foregoing aspects, the percentage similarity is at least about 60%, more preferably at least about 70%, still more preferably at least about 80%, even more preferably at least about 90% and yet even more preferably at least about 95% similar to at least about 5 contiguous amino acids, and more preferably at least about 10 contiguous amino acids of SEQ ID NO:4 or 5 or 6 or 7 or 19 and/or 20 or 22.

The present invention further extends to nucleic acid molecules, preferably in isolated form, encoding members of the CRP cytokine family. In one particular embodiment, the nucleic acid molecule comprises a sequence of nucleotides or a complementary sequence of nucleotides which encodes the amino acid sequence set forth in SEQ ID NO:4 or a sequence having at least 50% similarity thereto.

In another embodiment, the nucleic acid molecule comprises a sequence of nucleotides or a complementary sequence of nucleotides which encodes the amino acid sequence set forth in SEQ ID NO:5 or a sequence having at least 50% similarity thereto.

In a further embodiment, the nucleic acid molecule comprises a sequence of nucleotides or a complementary sequence of nucleotides which encodes the amino acid sequence set forth in SEQ ID NO:6 or a sequence having at least 50% similarity thereto.

In yet a further embodiment, the nucleic acid molecule comprises a sequence of nucleotides or a complementary sequence of nucleotides which encodes the amino acid sequence set forth in SEQ ID NO:7 or a sequence having at least 50% similarity thereto.

Still yet a further embodiment of the present invention provides a nucleic acid molecule

comprises a sequence of nucleotides or a complementary sequence of nucleotides which encodes the amino acid sequence set forth in SEQ ID NO:19 and/or SEQ ID NO:20 or a sequence having at least 50% similarity to either or both thereof.

Another embodiment of the present invention is directed to a nucleic acid molecule comprises a sequence of nucleotides or a complementary sequence of nucleotides which encodes the amino acid sequence set forth in SEQ ID NO:22 or a sequence having at least 50% similarity thereto.

In a particularly preferred embodiment, the nucleotide sequence is as set forth in SEQ ID NO:3 or a sequence having at least 50% similarity thereto and which is capable of hybridizing under low stringency conditions at 42°C to SEQ ID NO:3. The nucleotide sequence set forth in SEQ ID NO:3 encodes mCRP-1.

In another particularly preferred embodiment, the nucleotide sequence is as set forth in SEQ ID NO:18 or is a sequence having at least 50% similarity thereto and which is capable of hybridizing under low stringency conditions at 42°C to SEQ ID NO:18. The nucleotide sequence set forth in SEQ ID NO:18 is a genomic sequence encoding hCRP-1.

In yet another particularly preferred embodiment, the nucleotide sequence is as set forth in SEQ ID NO:21 or is a sequence having at least 50% similarity thereto and which is capable of hybridizing under low stringency conditions at 42°C to SEQ ID NO:21. The nucleotide sequence set forth in SEQ ID NO:18 encodes hCRP1 but without the intron.

Reference herein to a low stringency at 42°C includes and encompasses from at least about 1% v/v to at least about 15% v/v formamide and from at least about 1M to at least about 2M salt for hybridisation, and at least about 1M to at least about 2M salt for washing conditions. Alternative stringency conditions may be applied where necessary, such as medium stringency, which includes and encompasses from at least about 16% v/v to at least about 30% v/v formamide and from at least about 0.5M to at least about 0.9M salt for hybridisation, and at least about 0.5M to at least about 0.9M salt for washing conditions, or high stringency, which

includes and encompasses from at least about 31% v/v to at least about 50% v/v formamide and from at least about 0.01M to at least about 0.15M salt for hybridisation, and at least about 0.01M to at least about 0.15M salt for washing conditions.

Preferably, the percentage nucleotide similarity to SEQ ID NO:3 is at least about 60%, more preferably at least about 70%, still more preferably at least about 80%, even more preferably at least about 80% and yet even more preferably at least about 95% or above similarity to at least about 10 and more preferably at least about 20 contiguous nucleotides in SEQ ID NO:3.

Preferably, the percentage nucleotide similarity to SEQ ID NO:18 or SEQ ID NO:21 is at least about 60%, more preferably at least about 70%, still more preferably at least about 80%, even more preferably at least about 80% and yet even more preferably at least about 95% or above similarity to at least about 10 and more preferably at least about 20 contiguous nucleotides in SEQ ID NO:18 or SEQ ID NO:21.

The term "similarity" as used herein includes exact identity between compared sequences at the nucleotide or amino acid level. Where there is non-identity at the nucleotide level, "similarity" includes differences between sequences which result in different amino acids that are nevertheless related to each other at the structural, functional, biochemical and/or conformational levels. Where there is non-identity at the amino acid level, "similarity" includes amino acids that are nevertheless related to each other at the structural, functional, biochemical and/or conformational levels. The terms "homology" and "identity" as used herein can be considered equivalent to the term "similarity".

The CRP cytokines of the present invention are preferably but not exclusively of human, primate, laboratory test animal (eg. rabbit, guinea pig, mouse, rat), companion animal (eg. dog, cat), livestock animal (eg. sheep, cow, horse, donkey, pig) or captive wild animal (eg. deer, fox, kangaroo) origin.

The present invention encompasses a range of derivatives of the CRP cytokines. Derivatives include fragments, parts, portions, mutants, homologues and analogues of the CRP polypeptide

and corresponding genetic sequence. Derivatives also include single or multiple amino acid substitutions, deletions and/or additions to the CRP cytokine or single or multiple nucleotide substitutions, deletions and/or additions to the genetic sequence encoding the CRP cytokine. "Additions" to amino acid sequences or nucleotide sequences include fusions with other peptides, polypeptides or proteins or fusions to nucleotide sequences. Reference herein to the "CRP" cytokine includes reference to all derivatives thereof including functional and non-functional derivatives.

Analogues of the CRP cytokines contemplated herein include, but are not limited to, modification to side chains, incorporating of unnatural amino acids and/or their derivatives during peptide, polypeptide or protein synthesis and the use of crosslinkers and other methods which impose conformational constraints on the proteinaceous molecule or their analogues.

Examples of side chain modifications contemplated by the present invention include modifications of amino groups such as by reductive alkylation by reaction with an aldehyde followed by reduction with $NaBH_4$; amidination with methylacetimidate; acylation with acetic anhydride; carbamoylation of amino groups with cyanate; trinitrobenzylation of amino groups with 2, 4, 6-trinitrobenzene sulphonic acid (TNBS); acylation of amino groups with succinic anhydride and tetrahydrophthalic anhydride; and pyridoxylation of lysine with pyridoxal-5-phosphate followed by reduction with $NaBH_4$.

The guanidine group of arginine residues may be modified by the formation of heterocyclic condensation products with reagents such as 2,3-butanedione, phenylglyoxal and glyoxal.

The carboxyl group may be modified by carbodiimide activation *via* O-acylisourea formation followed by subsequent derivitisation, for example, to a corresponding amide.

Sulphydryl groups may be modified by methods such as carboxymethylation with iodoacetic acid or iodoacetamide; performic acid oxidation to cysteic acid; formation of a mixed disulphides with other thiol compounds; reaction with maleimide, maleic anhydride or other substituted maleimide; formation of mercurial derivatives using 4-chloromercuribenzoate, 4-

chloromercuriphenylsulphonic acid, phenylmercury chloride, 2-chloromercuri-4-nitrophenol and other mercurials; carbamoylation with cyanate at alkaline pH.

Tryptophan residues may be modified by, for example, oxidation with N-bromosuccinimide or alkylation of the indole ring with 2-hydroxy-5-nitrobenzyl bromide or sulphenyl halides. Tyrosine residues on the other hand, may be altered by nitration with tetranitromethane to form a 3-nitrotyrosine derivative.

Modification of the imidazole ring of a histidine residue may be accomplished by alkylation with iodoacetic acid derivatives or N-carbethoxylation with diethylpyrocarbonate.

Examples of incorporating unnatural amino acids and derivatives during peptide synthesis include, but are not limited to, use of norleucine, 4-amino butyric acid, 4-amino-3-hydroxy-5-phenylpentanoic acid, 6-aminohexanoic acid, t-butylglycine, norvaline, phenylglycine, ornithine, sarcosine, 4-amino-3-hydroxy-6-methylheptanoic acid, 2-thienyl alanine and/or D-isomers of amino acids. A list of unnatural amino acid, contemplated herein is shown in Table 1.

Crosslinkers can be used, for example, to stabilise 3D conformations, using homo-bifunctional crosslinkers such as the bifunctional imido esters having $(CH_2)_n$ spacer groups with n=1 to n=6, glutaraldehyde, N-hydroxysuccinimide esters and hetero-bifunctional reagents which usually contain an amino-reactive moiety such as N-hydroxysuccinimide and another group specific-reactive moiety such as maleimido or dithio moiety (SH) or carbodiimide (COOH). In addition, peptides can be conformationally constrained by, for example, incorporation of C_α and N_α-methylamino acids, introduction of double bonds between C_α and C_β atoms of amino acids and the formation of cyclic peptides or analogues by introducing covalent bonds such as forming an amide bond between the N and C termini, between two side chains or between a side chain and the N or C terminus.

The present invention further contemplates chemical analogues of the CRP cytokines capable of acting as antagonists or agonists of the CRP cytokine or which can act as functional

analogues of CRP cytokines. Chemical analogues may not necessarily be derived from the CRP cytokine but may share certain conformational similarities. Alternatively, chemical analogues may be specifically designed to mimic certain physiochemical properties of the CRP cytokine. Chemical analogues may be chemically synthesised or may be detected following, for example, natural product screening.

These types of modifications may be important to stabilise the CRP cytokine if administered to an individual or for use as a diagnostic reagent.

Other derivatives contemplated by the present invention include a range of glycosylation variants from a completely unglycosylated molecule to a modified glycosylated molecule. Altered glycosylation patterns may result from expression of recombinant molecules in different host cells.

TABLE 1

Non-conventional amino acid	Code	Non-conventional amino acid	Code
α-aminobutyric acid	Abu	L-N-methylalanine	Nmala
α-amino-α-methylbutyrate	Mgabu	L-N-methylarginine	Nmarg
aminocyclopropane-carboxylate	Cpro	L-N-methylasparagine	Nmasn
aminoisobutyric acid	Aib	L-N-methylaspartic acid	Nmasp
aminonorbornyl-carboxylate	Norb	L-N-methylcysteine	Nmcys
cyclohexylalanine	Chexa	L-N-methylglutamine	Nmgln
cyclopentylalanine	Cpen	L-N-methylglutamic acid	Nmglu
D-alanine	Dal	L-N-methylhistidine	Nmhis
D-arginine	Darg	L-N-methylisolleucine	Nmile
D-aspartic acid	Dasp	L-N-methylleucine	Nmleu
D-cysteine	Dcys	L-N-methyllysine	Nmlys
D-glutamine	Dgln	L-N-methylmethionine	Nmmet
D-glutamic acid	Dglu	L-N-methylnorleucine	Nmnle
D-histidine	Dhis	L-N-methylnorvaline	Nmnva
D-isoleucine	Dile	L-N-methylornithine	Nmorn
D-leucine	Dleu	L-N-methylphenylalanine	Nmphe
D-lysine	Dlys	L-N-methylproline	Nmpro
D-methionine	Dmet	L-N-methylserine	Nmser
D-ornithine	Dorn	L-N-methylthreonine	Nmthr
D-phenylalanine	Dphe	L-N-methyltryptophan	Nmtrp
D-proline	Dpro	L-N-methyltyrosine	Nmtyr
D-serine	Dser	L-N-methylvaline	Nmval
D-threonine	Dthr	L-N-methylethylglycine	Nmetg
D-tryptophan	Dtrp	L-N-methyl-t-butylglycine	Nmtbug
		L-norleucine	Nle
		L-norvaline	Nva

D-tyrosine	Dtyr	α-methyl-aminoisobutyrate	Maib
D-valine	Dval	α-methyl-γ-aminobutyrate	Mgabu
D-α-methylalanine	Dmala	α-methylcyclohexylalanine	Mchexa
D-α-methylarginine	Dmarg	α-methylcylcopentylalanine	Mcpen
D-α-methylasparagine	Dmasn	α-methyl-α-napthylalanine	Manap
D-α-methylaspartate	Dmasp	α-methylpenicillamine	Mpen
D-α-methylcysteine	Dmcys	N-(4-aminobutyl)glycine	Nglu
D-α-methylglutamine	Dmgln	N-(2-aminoethyl)glycine	Naeg
D-α-methylhistidine	Dmhis	N-(3-aminopropyl)glycine	Norn
D-α-methylisoleucine	Dmile	N-amino-α-methylbutyrate	Nmaabu
D-α-methylleucine	Dmleu	α-napthylalanine	Anap
D-α-methyllysine	Dmlys	N-benzylglycine	Nphe
D-α-methylmethionine	Dmmet	N-(2-carbamylethyl)glycine	Ngln
D-α-methylornithine	Dmorn	N-(carbamylmethyl)glycine	Nasn
D-α-methylphenylalanine	Dmphe	N-(2-carboxyethyl)glycine	Nglu
D-α-methylproline	Dmpro	N-(carboxymethyl)glycine	Nasp
D-α-methylserine	Dmser	N-cyclobutylglycine	Ncbut
D-α-methylthreonine	Dmthr	N-cycloheptylglycine	Nchep
D-α-methyltryptophan	Dmtrp	N-cyclohexylglycine	Nchex
D-α-methyltyrosine	Dmty	N-cyclodecylglycine	Ncdec
D-α-methylvaline	Dmval	N-cylcododecylglycine	Ncdod
D-N-methylalanine	Dnmala	N-cyclooctylglycine	Ncoct
D-N-methylarginine	Dnmarg	N-cyclopropylglycine	Ncpro
D-N-methylasparagine	Dnmasn	N-cycloundecylglycine	Ncund
D-N-methylaspartate	Dnmasp	N-(2,2-diphenylethyl)glycine	Nbhm
D-N-methylcysteine	Dnmcys	N-(3,3-diphenylpropyl)glycine	Nbhe
D-N-methylglutamine	Dnmgln	N-(3-guanidinopropyl)glycine	Narg
D-N-methylglutamate	Dnmglu	N-(1-hydroxyethyl)glycine	Nthr
D-N-methylhistidine	Dnmhis	N-(hydroxyethyl))glycine	Nser
D-N-methylisoleucine	Dnmile	N-(imidazolylethyl))glycine	Nhis
D-N-methylleucine	Dnmleu	N-(3-indolylyethyl)glycine	Nhtrp

D-N-methyllysine	Dnmlys	N-methyl-γ-aminobutyrate	Nmgabu
N-methylcyclohexylalanine	Nmchexa	D-N-methylmethionine	Dnmmet
D-N-methylornithine	Dnmorn	N-methylcyclopentylalanine	Nmcpen
N-methylglycine	Nala	D-N-methylphenylalanine	Dnmphe
N-methylaminoisobutyrate	Nmaib	D-N-methylproline	Dnmpro
N-(1-methylpropyl)glycine	Nile	D-N-methylserine	Dnmser
N-(2-methylpropyl)glycine	Nleu	D-N-methylthreonine	Dnmthr
D-N-methyltryptophan	Dnmtrp	N-(1-methylethyl)glycine	Nval
D-N-methyltyrosine	Dnmtyr	N-methyla-napthylalanine	Nmanap
D-N-methylvaline	Dnmval	N-methylpenicillamine	Nmpen
γ-aminobutyric acid	Gabu	N-(*p*-hydroxyphenyl)glycine	Nhtyr
L-*t*-butylglycine	Tbug	N-(thiomethyl)glycine	Ncys
L-ethylglycine	Etg	penicillamine	Pen
L-homophenylalanine	Hphe	L-α-methylalanine	Mala
L-α-methylarginine	Marg	L-α-methylasparagine	Masn
L-α-methylaspartate	Masp	L-α-methyl-*t*-butylglycine	Mtbug
L-α-methylcysteine	Mcys	L-methylethylglycine	Metg
L-α-methylglutamine	Mgln	L-α-methylglutamate	Mglu
L-α-methylhistidine	Mhis	L-α-methylhomophenylalanine	Mhphe
L-α-methylisoleucine	Mile	N-(2-methylthioethyl)glycine	Nmet
L-α-methylleucine	Mleu	L-α-methyllysine	Mlys
L-α-methylmethionine	Mmet	L-α-methylnorleucine	Mnle
L-α-methylnorvaline	Mnva	L-α-methylornithine	Morn
L-α-methylphenylalanine	Mphe	L-α-methylproline	Mpro
L-α-methylserine	Mser	L-α-methylthreonine	Mthr
L-α-methyltryptophan	Mtrp	L-α-methyltyrosine	Mtyr
L-α-methylvaline	Mval	L-N-methylhomophenylalanine	Nmhphe

N-(N-(2,2-diphenylethyl) carbamylmethyl)glycine	Nnbhm	N-(N-(3,3-diphenylpropyl) carbamylmethyl)glycine	Nnbhe
1-carboxy-1-(2,2-diphenyl-ethylamino)cyclopropane	Nmbc		

The identification of a new family of CRP cytokines allows definition of a consensus protein motif which may be used to identify further family members by, for example, PCR analysis of cDNA and/or genomic libraries.

The identification of the CRP cytokines of the present invention also permits the generation of a range of therapeutic molecules capable of modulating expression of the CRP cytokine or modulating the activity of the cytokine. Modulators contemplated by the present invention includes agonists and antagonists of CRP cytokine expression. Antagonists of CRP cytokine expression include antisense molecules, ribozymes and co-suppression molecules. Agonists include molecules which increase promoter activity or interfere with negative regulatory mechanisms. Antagonists of the cytokine include antibodies, inhibitor peptide fragments and soluble receptors.

Another embodiment of the present invention contemplates a method for modulating expression of a CRP cytokine in a mammal such as a human, primate or laboratory test animal, said method comprising contacting a gene encoding said CRP cytokine with an effective amount of a modulator of CRP cytokine expression for a time and under conditions sufficient to up-regulate or down-regulate or otherwise modulate expression of the CRP cytokine. For example, a nucleic acid molecule encoding the CRP cytokine or a derivative thereof may be introduced into a cell to enhance the ability of that cell to produce CRP or, through co-suppression reduce CRP gene expression. Alternatively, CRP cytokine antisense sequences such as oligonucleotides may be introduced to decrease CRP expression in any cell expressing the endogenous CRP cytokine gene.

Another aspect of the present invention contemplates a method of modulating activity of the

CRP cytokine in a mammalian such as a human, said method comprising administering to said mammal a modulating effective amount of a molecule for a time and under conditions sufficient to increase or decrease CRP cytokine activity. The molecule may be a proteinaceous molecule or a chemical entity and may also be a derivative of a CRP cytokine or its ligand (eg. a receptor) or a chemical analogue or truncation mutant of a CRP cytokine or its ligand.

Modulating CRPs may be important in a number of respects, such as but not limited to modulating early embryo development. For example, CRPs may have a role in somite patterning and in muscles of the body and limbs. The function of the CRPs may be in the muscles themselves (autocrine role), or may be to act on neighbouring tissue of the somite, the dermatome (forming skin) and schlerotome (forming bone). CRP may be involved in craniofacial development and useful clinically as an inductive, maintenance, survival, proliferative, anti-proliferative or differentiation factor in pathologies related to muscle, bone and skin. Furthermore, it may act as a tumour suppressor suggesting a function as an antiproliferative factor in a broad range of cancers, including breast cancer, lymphoma and leukaemia, melanoma, colorectal cancer, pancreatic cancer, lung cancer, stomach cancer and neuroblastoma.

The CRPs are also good candidates for a inductive factor that either induce or give anterior character to, early neural tissue. Such as they may be clinically useful as an inductive, maintenance, survival, proliferative, anti-proliferative or differentiation factor in any degenerative neuropathy or in any grafting procedure involving foetal or neural tissue grafted to correct familial or acquired deficiencies, or to repair neural tissue after trauma. Anterior lateral endoderm is also known to be the source of inducers and suppressors of cardiogenesis and the CRPs may comprise these factors. They may also be useful clinically as an inductive, maintenance, survival, proliferative, anti-proliferative or differentiation factor for cadiomyocytes in any interventionist procedure to correct cardiomypathgy, including any grafting technique aimed at repairing infarcts (cardiomyoplasty) or valvular defects, modification (including inhibition) or familial or secondary hypertrophy, or induction of compensatory cardiomycyte growth during ageing.

Accordingly, the present invention contemplates a pharmaceutical composition comprising one or more CRP cytokines or derivatives thereof or a modulator of CRP cytokine expression or CRP cytokine activity and one or more pharmaceutically acceptable carriers and/or diluents. These components are referred to as the active ingredients.

The pharmaceutical forms suitable for injectable use include sterile aqueous solutions (where water soluble) and sterile powders for the extemporaneous preparation of sterile injectable solutions. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol and liquid polyethylene glycol, and the like), suitable mixtures thereof and vegetable oils. The proper fluidity can be maintained, for example, by the use of a coating such as licithin. The preventions of the action of microorganisms can be brought about by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thimerosal and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars or sodium chloride. Prolonged absorption of the injectable compositions can be brought about by the use in the compositions of agents delaying absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions are prepared by incorporating the active compounds in the required amount in the appropriate solvent with various of the other ingredients enumerated above, as required, followed by filtered sterilization. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and the freeze-drying technique which yield a powder of the active ingredient plus any additional desired ingredient from previously sterile-filtered solution thereof.

When the active ingredients are suitably protected they may be orally administered, for example, with an inert diluent or with an assimilable edible carrier, or it may be enclosed in hard or soft shell gelatin capsule, or it may be compressed into tablets, or it may be incorporated directly with the food of the diet. For oral therapeutic administration, the active compound may be incorporated with excipients and used in the form of ingestible tablets,

buccal tablets, troches, capsules, elixirs, suspensions, syrups, wafers, and the like. Such compositions and preparations should contain at least 1% by weight of active compound. The percentage of the compositions and preparations may, of course, be varied and may conveniently be between about 5 to about 80% of the weight of the unit. The amount of active compound in such therapeutically useful compositions in such that a suitable dosage will be obtained. Preferred compositions or preparations according to the present invention are prepared so that an oral dosage unit form contains between about 0.01 μg and about 2000 mg of active compound. Alternative amounts include between about 1.0 μg and about 1500 ng, between about 1μg and about 1000 mg and between about 10 μg and about 500 mg.

The tablets, troches, pills, capsules and the like may also contain the components as listed hereafter: A binder such as gum, acacia, corn starch or gelatin; excipients such as dicalcium phosphate; a disintegrating agent such as corn starch, potato starch, alginic acid and the like; a lubricant such as magnesium stearate; and a sweetening agent such a sucrose, lactose or saccharin may be added or a flavouring agent such as peppermint, oil of wintergreen, or cherry flavouring. When the dosage unit form is a capsule, it may contain, in addition to materials of the above type, a liquid carrier. Various other materials may be present as coatings or to otherwise modify the physical form of the dosage unit. For instance, tablets, pills, or capsules may be coated with shellac, sugar or both. A syrup or elixir may contain the active compound, sucrose as a sweetening agent, methyl and propylparabens as preservatives, a dye and flavouring such as cherry or orange flavour. Of course, any material used in preparing any dosage unit form should be pharmaceutically pure and substantially non-toxic in the amounts employed. In addition, the active compound(s) may be incorporated into sustained-release preparations and formulations.

The present invention also extends to forms suitable for topical application such as creams, lotions and gels.

Pharmaceutically acceptable carriers and/or diluents include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents and the like. The use of such media and agents for pharmaceutical active substances is well

known in the art. Except insofar as any conventional media or agent is incompatible with the active ingredient, use thereof in the therapeutic compositions is contemplated. Supplementary active ingredients can also be incorporated into the compositions.

It is especially advantageous to formulate parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the mammalian subjects to be treated; each unit containing a predetermined quantity of active material calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the novel dosage unit forms of the invention are dictated by and directly dependent on (a) the unique characteristics of the active material and the particular therapeutic effect to be achieved, and (b) the limitations inherent in the art of compounding such an active material for the treatment of disease in living subjects having a diseased condition in which bodily health is impaired as herein disclosed in detail.

The principal active ingredient is compounded for convenient and effective administration in effective amounts with a suitable pharmaceutically acceptable carrier in dosage unit form. A unit dosage form can, for example, contain the principal active compound in amounts ranging from 0.01 µg to about 2000 mg. Expressed in proportions, the active compound is generally present in from about 0.5 µg to about 2000 mg/ml of carrier. In the case of compositions containing supplementary active ingredients, the dosages are determined by reference to the usual dose and manner of administration of the said ingredients. Alternatively, amounts administered may be represented in terms of amounts/kg body weight. In this case, amounts range from about 0.001 μg to about 1000 mg/kg body weight may be administered 500 mg/kg body weight or about 10.01 μg to about or above 0.1 μg to about 250 mg/kg body weight are contemplated by the present invention.

The pharmaceutical composition may also comprise genetic molecules such as a vector capable of transfecting target cells where the vector carries a nucleic acid molecule capable of modulating CRP expression or CRP activity. The vector may, for example, be a viral vector.

Still another aspect of the present invention is directed to antibodies to CRP cytokines and their derivatives. Such antibodies may be monoclonal or polyclonal and may be selected from naturally occurring antibodies to CRP cytokines or may be specifically raised to CRP cytokines or derivatives thereof. In the case of the latter, a CRP cytokine or its derivative may first need to be associated with a carrier molecule. The antibodies and/or recombinant CRP cytokine or its derivatives of the present invention are particularly useful as therapeutic or diagnostic agents.

For example, a CRP cytokine and its derivatives can be used to screen for naturally occurring antibodies to CRP cytokines. These may occur, for example in some autoimmune diseases. Alternatively, specific antibodies can be used to screen for CRP cytokines. Techniques for such assays are well known in the art and include, for example, sandwich assays and ELISA. Knowledge of CRP cytokines levels may be important for diagnosis of certain cancers or a predisposition to cancers or for monitoring certain therapeutic protocols.

Antibodies to CRP cytokines of the present invention may be monoclonal or polyclonal. Alternatively, fragments of antibodies may be used such as Fab fragments. Furthermore, the present invention extends to recombinant and synthetic antibodies and to antibody hybrids. A "synthetic antibody" is considered herein to include fragments and hybrids of antibodies. The antibodies of this aspect of the present invention are particularly useful for immunotherapy and may also be used as a diagnostic tool for assessing apoptosis, cancer, tissue regeneration or development, muscle development or health of neural tissue or for monitoring the program of a therapeutic regimin.

For example, specific antibodies can be used to screen for CRP proteins. The latter would be important, for example, as a means for screening for levels of a CRP in a cell extract or other biological fluid or purifying a CRP made by recombinant means from culture supernatant fluid. Techniques for the assays contemplated herein are known in the art and include, for example, sandwich assays and ELISA.

It is within the scope of this invention to include any second antibodies (monoclonal,

polyclonal or fragments of antibodies or synthetic antibodies) directed to the first mentioned antibodies discussed above. Both the first and second antibodies may be used in detection assays or a first antibody may be used with a commercially available anti-immunoglobulin antibody. An antibody as contemplated herein includes any antibody specific to any region of CRP.

Both polyclonal and monoclonal antibodies are obtainable by immunization with the enzyme or protein and either type is utilizable for immunoassays. The methods of obtaining both types of sera are well known in the art. Polyclonal sera are less preferred but are relatively easily prepared by injection of a suitable laboratory animal with an effective amount of a CRP, or antigenic parts thereof, collecting serum from the animal, and isolating specific sera by any of the known immunoadsorbent techniques. Although antibodies produced by this method are utilizable in virtually any type of immunoassay, they are generally less favoured because of the potential heterogeneity of the product.

The use of monoclonal antibodies in an immunoassay is particularly preferred because of the ability to produce them in large quantities and the homogeneity of the product. The preparation of hybridoma cell lines for monoclonal antibody production derived by fusing an immortal cell line and lymphocytes sensitized against the immunogenic preparation can be done by techniques which are well known to those who are skilled in the art.

Another aspect of the present invention contemplates a method for detecting CRP in a biological sample from a subject said method comprising contacting said biological sample with an antibody specific for a CRP(or group of CRP s) or its derivatives or homologues for a time and under conditions sufficient for an antibody-CRP complex to form, and then detecting said complex.

The presence of CRP may be accomplished in a number of ways such as by Western blotting and ELISA procedures. A wide range of immunoassay techniques are available as can be seen by reference to US Patent Nos. 4,016,043, 4, 424,279 and 4,018,653. These, of course, includes both single-site and two-site or "sandwich" assays of the non-competitive types, as

well as in the traditional competitive binding assays. These assays also include direct binding of a labelled antibody to a target.

Sandwich assays are among the most useful and commonly used assays and are favoured for use in the present invention. A number of variations of the sandwich assay technique exist, and all are intended to be encompassed by the present invention. Briefly, in a typical forward assay, an unlabelled antibody is immobilized on a solid substrate and the sample to be tested brought into contact with the bound molecule. After a suitable period of incubation, for a period of time sufficient to allow formation of an antibody-antigen complex, a second antibody specific to the antigen, labelled with a reporter molecule capable of producing a detectable signal is then added and incubated, allowing time sufficient for the formation of another complex of antibody-antigen-labelled antibody. Any unreacted material is washed away, and the presence of the antigen is determined by observation of a signal produced by the reporter molecule. The results may either be qualitative, by simple observation of the visible signal, or may be quantitated by comparing with a control sample containing known amounts of hapten. Variations on the forward assay include a simultaneous assay, in which both sample and labelled antibody are added simultaneously to the bound antibody. These techniques are well known to those skilled in the art, including any minor variations as will be readily apparent. In accordance with the present invention the sample is one which might contain CRP including cell extract, tissue biopsy or possibly serum, saliva, mucosal secretions, lymph, tissue fluid and respiratory fluid. The sample is, therefore, generally a biological sample comprising biological fluid but also extends to fermentation fluid and supernatant fluid such as from a cell culture.

In the typical forward sandwich assay, a first antibody having specificity for the CRP or antigenic parts thereof, is either covalently or passively bound to a solid surface. The solid surface is typically glass or a polymer, the most commonly used polymers being cellulose, polyacrylamide, nylon, polystyrene, polyvinyl chloride or polypropylene. The solid supports may be in the form of tubes, beads, discs of microplates, or any other surface suitable for conducting an immunoassay. The binding processes are well-known in the art and generally consist of cross-linking covalently binding or physically adsorbing, the polymer-antibody

complex is washed in preparation for the test sample. An aliquot of the sample to be tested is then added to the solid phase complex and incubated for a period of time sufficient (e.g. 2-40 minutes or where more convenient, overnight) and under suitable conditions (e.g. for about 20°C to about 40°C) to allow binding of any subunit present in the antibody. Following the incubation period, the antibody subunit solid phase is washed and dried and incubated with a second antibody specific for a portion of the hapten. The second antibody is linked to a reporter molecule which is used to indicate the binding of the second antibody to the hapten.

An alternative method involves immobilizing the target molecules in the biological sample and then exposing the immobilized target to specific antibody which may or may not be labelled with a reporter molecule. Depending on the amount of target and the strength of the reporter molecule signal, a bound target may be detectable by direct labelling with the antibody. Alternatively, a second labelled antibody, specific to the first antibody is exposed to the target-first antibody complex to form a target-first antibody-second antibody tertiary complex. The complex is detected by the signal emitted by the reporter molecule.

By "reporter molecule" as used in the present specification, is meant a molecule which, by its chemical nature, provides an analytically identifiable signal which allows the detection of antigen-bound antibody. Detection may be either qualitative or quantitative. The most commonly used reporter molecules in this type of assay are either enzymes, fluorophores or radionuclide containing molecules (i.e. radioisotopes) and chemiluminescent molecules.

In the case of an enzyme immunoassay, an enzyme is conjugated to the second antibody, generally by means of glutaraldehyde or periodate. As will be readily recognized, however, a wide variety of different conjugation techniques exist, which are readily available to the skilled artisan. Commonly used enzymes include horseradish peroxidase, glucose oxidase, beta-galactosidase and alkaline phosphatase, amongst others. The substrates to be used with the specific enzymes are generally chosen for the production, upon hydrolysis by the corresponding enzyme, of a detectable colour change. Examples of suitable enzymes include alkaline phosphatase and peroxidase. It is also possible to employ fluorogenic substrates, which yield a fluorescent product rather than the chromogenic substrates noted above. In all cases, the enzyme-labelled antibody is added to the first antibody hapten complex, allowed to

bind, and then the excess reagent is washed away. A solution containing the appropriate substrate is then added to the complex of antibody-antigen-antibody. The substrate will react with the enzyme linked to the second antibody, giving a qualitative visual signal, which may be further quantitated, usually spectrophotometrically, to give an indication of the amount of hapten which was present in the sample. "Reporter molecule" also extends to use of cell agglutination or inhibition of agglutination such as red blood cells on latex beads, and the like.

Alternately, fluorescent compounds, such as fluorescein and rhodamine, may be chemically coupled to antibodies without altering their binding capacity. When activated by illumination with light of a particular wavelength, the fluorochrome-labelled antibody adsorbs the light energy, inducing a state to excitability in the molecule, followed by emission of the light at a characteristic colour visually detectable with a light microscope. As in the EIA, the fluorescent labelled antibody is allowed to bind to the first antibody-hapten complex. After washing off the unbound reagent, the remaining tertiary complex is then exposed to the light of the appropriate wavelength the fluorescence observed indicates the presence of the hapten of interest. Immunofluorescene and EIA techniques are both very well established in the art and are particularly preferred for the present method. However, other reporter molecules, such as radioisotope, chemiluminescent or bioluminescent molecules, may also be employed.

Although many variations exist for such immunologically based assays another alternative is to identify and isolate the cytokine receptor and immobilize this to a solid support in such a way that it can still bind the cytokine. A sample containing the cytokine is brought into contact with the immobilized receptor to permit the receptor cytokine complex to form. Bound cytokine is then identified using an antibody to the cytokine. The antibody may be labelled or detected using another anti-immunoglobulin antibody which is labelled.

The present invention also contemplates genetic assays such as involving PCR analysis to detect a CRP gene or its derivatives. Alternative methods or methods used in conjunction include direct nucleotide sequencing or mutation scanning such as single stranded conformation polymorphoms analysis (SSCP) as specific oligonucleotide hybridisation, as methods such as direct protein truncation tests.

The nucleic acid molecules of the present invention may be RNA or DNA. When the nucleic acid molecule is in DNA form, it may be genomic DNA or cDNA. RNA forms of the nucleic acid molecules of the present invention are generally mRNA.

Although the nucleic acid molecules of the present invention are generally in isolated form, they may be integrated into or ligated to or otherwise fused or associated with other genetic molecules such as vector molecules and in particular expression vector molecules. Vectors and expression vectors are generally capable of replication and, if applicable, expression in one or both of a prokaryotic cell or a eukaryotic cell. Preferably, prokaryotic cells include *E. coli*, *Bacillus sp* and *Pseudomonas sp*. Preferred eukaryotic cells include yeast, fungal, mammalian and insect cells.

Accordingly, another aspect of the present invention contemplates a genetic construct comprising a vector portion and a mammalian such as a human CRP gene portion, which CRP gene portion is capable of encoding a CRP polypeptide or a functional or immunologically interactive derivative thereof.

Preferably, the CRP gene portion of the genetic construct is operably linked to a promoter on the vector such that said promoter is capable of directing expression of said CRP gene portion in an appropriate cell.

In addition, the CRP gene portion of the genetic construct may comprise all or part of the gene fused to another genetic sequence such as a nucleotide sequence encoding glutathione-S-transferase or part thereof. It is also within the scope of the present invention to include fusions between CRP cytokines. Such fusion molecules may have increased pleiotropy and/or provides a means of, for example, "humanising" a non-human CRP.

The present invention extends to such genetic constructs and to prokaryotic or eukaryotic cells comprising same.

The present invention also extends to any or all derivatives of CRP cytokines including

mutants, part, fragments, portions, homologues and analogues or their encoding genetic sequence including single or multiple nucleotide or amino acid substitutions, additions and/or deletions to the naturally occurring nucleotide or amino acid sequence.

The CRP cytokines and their genetic sequences of the present invention will be useful in the generation of a range of therapeutic and diagnostic reagents and will be especially useful in the detection of ligands (eg. receptors) capable of interacting with CRP cytokines. For example, a recombinant CRP may be bound or fused to a reporter molecule capable of producing an identifiable signal, contacted with a cell or group of cells putatively carrying a CRP ligand (eg. a receptor) and screening for binding of the labelled CRP to a ligand. Alternatively, labelled CRP may be used to screen expression libraries of putative ligands.

CRP cytokines are important for the regulation, maintenance and survival of a diverse array of cell types such as but not limited to muscle tissue, bone, skin and/or neural tissue. Accordingly, it is proposed that CRP cytokines or their functional derivatives be used to regulate development, maintenance or regeneration in an array of different cells and tissues *in vitro* and *in vivo*. For example, CRPs are contemplated to be useful in modulating neuronal proliferation, differentation and survival.

Soluble CRP polypeptides are also contemplated to be useful in the treatment of disease, injury or abnormality in the nervous system, e.g. in relation to central or peripheral nervous system to treat Cerebral Palsy, trauma induced paralysis, vascular ischaemia associated with stroke, neuronal tumours, motoneurone disease, Parkinson's disease, Huntington's disease, Alzheimer's disease, Multiple Sclerosis, peripheral neuropathies associated with diabetes, heavy metal or alcohol toxicity, renal failure and infectious diseases such as herpes, rubella, measles, chicken pox, HIV or HTLV-1.

A membrane bound CRP may be used *in vitro* on nerve cells or tissues to modulate proliferation, differentiation or survival, for example, in grafting procedures or transplantation.

As stated above, the CRP cytokines of the present invention or their functional derivatives may

be provided in a pharmaceutical composition together with one or more pharmaceutically acceptable carriers and/or diluents. In addition, the present invention contemplates a method of treatment comprising the administration of an effective amount of a CRP of the present invention. The present invention also extends to antagonists and agonists of CRP molecules and their use in therapeutic compositions and methodologies.

SUMMARY OF SEQ ID NOs

Sequence	SEQ ID NO.
Cystine knot region amino acid sequences	1,2
nucleotide sequence of mCRP-1	3
amino acid sequence of mCRP-1	4
amino acid sequence of rCRP-2	5
amino acid sequence of mCRP-2	6
amino acid sequence of hCRP-2	7
oligonucleotide primers	8-16
FLAG epitope	17
genomic nucleotide sequence of hCRP-1	18
part amino acid sequence of genomic hCRP-1	19
part amino acid sequence of genomic hCRP-1	20
coding sequence of hCRP-1 without intron	21
amino acid sequence of SEQ ID NO:20	22
N-terminal amino acid sequence of mCRP-1	23
N-terminal amino acid sequence of mCRP-2	24

The following single and three letter abbreviations are used for amino acid residues:

Amino Acid	Three-letter Abbreviation	One-letter Symbol
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamine	Gln	Q
Glutamic acid	Glu	E
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V
Any residue	Xaa	X

EXAMPLE 1

Figure 1 is an alignment of amino acid sequences comparing *Xenopus laevis* cerberus and members of the CRP family from mouse, rat and human sources. CRP proteins share approximately 30% identity and 40% similarity over an 88 amino acid region. mCRP-1 and mCer-1 are used interchangeably in the specification to refer to the same molecule. Xcer refers to cerberus.

EXAMPLE 2

CRP is a secretable molecule

To examine whether CRP is secreted, mouse CRP-2 (mCRP-2) with a FLAG epitope fused to its C-terminus was transiently expressed in COS cells. A single epitope-tagged protein of 29kDa was secreted from mCRP-2-transfected cells (Figure 4), but not from those transfected with empty FLAG vector. Indicative of passage through the rough ER and golgi, secreted CRP was N-glycosylated (Figure 4). Secreted mCRP-2 ran at approximately 30 kDa after deglycosylation. N-terminal amino acid sequence determination confirmed the identity of secreted mCRP-2 and the fact that the hydrophobic signal peptide had been processed (see Figure 1).

EXAMPLE 3

Isolation of a full length cDNA encoding a mouse CRP

A full length cDNA encoding a mouse CRP (mCRP-1) was isolated using oligonucleotides based on the Genbank EST AA120122, a mouse embryonic day 7.5 cDNA, in combination with oligonucleotides corresponding to vector sequences of the plasmid pSport-1. The nucleotide and corresponding amino acid sequence is shown in Figure 2. A comparison of the mCRP-1 and cerberus sequence is shown in Figure 1.

An amount of 50 ng of mouse embryonic region cDNA library was used as template in PCR which contained two primers.

Reaction 1: Oligo 1 (CTTGGAAGATTCTGGAAGAAACCTG [SEQ ID NO:8]) X m13

forward (CGCCAGGGTTTTCCCAGTCACGAC [SEQ ID NO:9])

Reaction 2: Oligo 3 (GCCCCTTCTCCGGGAAAACGAATG [SEQ ID NO:10]) X m13 reverse (GGAAACAGCTATGACCATGATTAC [SEQ ID NO:11])

The PCR was performed using 1 unit of Promega Taq polymerase in the buffer supplied in the presence of 200μ M dNTPs and 2.5 mM $MgCl_2$ for 30 cycles at [94°C, 60°C, 73°C] and 10, 30, 60 seconds]. The products of these reactions were diluted 1 in 100 in water and a second PCR was then performed on the diluted sample.

Reaction 3 (template = diluted reaction 1) Oligo 2 (CAGGACTGTGCCCTTCAACCAGAC [SEQ ID NO:13]) X Atttohind (GACGTCGCATGCACGCGTACGTAAG [SEQ ID NO:12])

Reaction 4 (template = diluted reaction 2) Oligo 4 (CGGTCTCAGGTTTCTTCCAGAATC [SEQ ID NO:13]) X PsttoEco (CTGCAGGTACCGGTCCGGAATTCC [SEQ ID NO:14])

All PCRs contained 50 ng of each primer.

Each of these reactions yielded a prominant product of approximately 500 base pairs. The products were gel purified using Bresaclean fragment purification system and cloned into the *Hinc*II site of pBluscript SK+(Stratagene). Recombinant plasmids were purified and their inserts sequenced using the m13 forward and reverse primers described above in standard ABI sequencing reactions. The sequence data obtained from these were combined with the EST data to construct a continuous 1000 base pair sequence from which the putative full length protein homologous to *Xenopus* cerberus could be translated. Two further primers Oligo 6 (CCATCTGTGAATCTAACCTCAGTCTC [SEQ ID NO:15] and Oligo 7 (AACTCACATAACATTTCCAGATTG [SEQ ID NO:16]) lying at the extremities of this sequence were used in PCR under the same conditions as above (with the day 7.5 embryo library as template) to generate a DNA fragment completely encompassing the putative coding sequence.

EXAMPLE 4

Production and analysis of mouse CRP-2 protein

Protein Production

For analysis of structure and functional activity, mouse CRP (mCRP-2) with a FLAG epitope (DYKDDDDK [SEQ ID NO:18]) fused to its C-terminus was transiently expressed in COS cells. COS cells were transfected with the mCRP-2 cDNA using a polycationic liposome transfection reagent (Lipofetamine, GibcoBRL) or electroporation (Gene Pulser, Biorad). For lipofectamine mediated transfection, COS cells grown to approximately 70-80% confluence in 100 mm petri-dishes were washed in serum free DMEM media then exposed to a mixture of mCRP cDNA and lipofectamine diluted in DMEM according to the manufacturers instructions. After 5 hrs incubation at 37°C with 5%v/v CO_2 the cells were washed once with DMEM and incubated for a further 16 hrs in DMEM supplemented with 10% v/v FCS, glutamine and antibiotics (DM10). At this time the DM10 was removed and replaced with a further 8 ml/dish of fresh DM10 and transfected cells incubated from a further 48 hrs. Supernatents containing secreted mouse CRP were recovered, centrifuged and filtered to remove cell debris, then stored at 4°C. For electroporation 1 x 10^7 COS cells were resuspended in 400 μl of DMEM and transferred to a 0.4 cm electroporation cuvette containing 20 μg of mCRP cDNA and incubated for 10 min at room temperature. After electroporation (0.3 kV; 960 μFD) cells were incubated at room temperature for a further 10 min then transferred to a 150 cm^2 tissue culture flask containing 40 ml of DM10. Supernatant containing secreted mCRP-2 protein was recovered after 72 hrs, centrifuged and filtered to remove cell debris, then stored at 4°C.

Protein Analysis

Expression of mCRP-2 was monitored by biosensor analysis and western blot analysis. Samples were monitored for expression using a biosensor (BIAcore™) employing surface plasmon resonance detection. M2 antibody, specific to FLAG sequence, was immobilised to the sensor surface using NHS-EDC coupling chemistry according to the manufacturers instructions. The running buffer for all biosensor analysis was 10 mM HEPES, 150mM NaCl, 3.4mM EDTA, 0.005% v/v Tween 20 (HBS buffer). The buffer was degassed under vacuum

for 10 minutes prior to use to prevent bubble formation. The flow rate was 5 μl/min. For immobilisation of M2 antibody, 50mM N-hydroxysuccinimide and 200mM N-ethyl-'dimethylaminopropyl carbodiimide were mixed in a 1:1 ratio and 35 μl of this mixture injected onto the sensorchip for surface activation. After 7 minutes, 35 μl of M2 antibody (100 μg/ml in 20mM Sodium Acetate pH 4.2) was injected. After a further 7 minutes, 75 μl ethanolamine was injected to quench any remaining free esters generated during the NHS-EDC activation.

For analysis of binding, 35 μl of sample supernatant was injected onto the sensor surface and allowed to bind to the M2 derivitised channel. The differential between the signal prior to injection and post injection (450 seconds) was recorded as the specific response units (resonance units). Post sample analysis, 10 μl of 50mM diethylamine pH 12.0 was used to desorb the sensor surface prior to the subsequent analysis. mCRP-2 expression was monitored on days 1, 2 and 3 post-transfection and compared to a control supernatant generated by transfection of a cDNA encoding an unrelated protein (β-galactosidase). Biosensor results are shown in Figure 4.

Protein expression was confirmed by western blot analysis. Samples were separated on the basis of size using SDS-PAGE analysis. Samples were then blotted onto PVDF using Tris/glycine/Methanol buffer at 30 V for 1 hour using a Novex blot apparatus. Transfers were performed according to manufacturers instructions. PVDF membranes were then blocked overnight in 2% w/v BSA in TBS buffer (20mM Tris pH 7.4, 150 mM NaCl, 0.02 v/v Tween 20). Membranes were washed extensively between addition of antibodies with 10 washes of TBS. Blots were probed with primary antibody (M2) for 1hr at a final concentration of 10 μg/ml. Bound M2 was detected using an appropriately conjugated secondary antibody (goat anti-mouse HRP, Silenus) at a dilution of 1:5000 for 1 hour. Bound secondary antibody was visualised by autoradiography after addition of ECL reagent.

Protein Purification

Proteins were purified using the strategy outline in Figure 5. Binding was monitored on the Biosensor and elution performed, after extensive washing in TBS (200 column volumes),

using 4 x 5 ml fractions of Flag peptide (60 μg/ml) in TBS. Fractions were analysed by SDS-PAGE analysis under reducing conditions using the PHAST system (Pharmacia) and protein visualised by silver staining and western blot analysis.

Protein Characterisation

Using SDS-PAGE under non reducing and reducing conditions (Figure 6) mCRP-2 was examined to determine the size of the expressed protein. Under non-reducing conditions a 65 kD protein was observed, whereas under reducing conditions mCRP-2 had an apparent molecular weight of 30 kD. These data suggest that the bulk of the mCRP-2 is expressed as a disulphide bonded homodimer. This observation was confirmed by size exclusion analysis of CRP using a Superose 12 PC 3.2/30 column. Column fractions containing mCRP were detected by biosensor analysis and the size of the expressed protein determined by linear regression analysis (Figure 7).

EXAMPLE 5

Production and analysis of mouse CRP-1 protein

Protein Production

For analysis of structure and functional activity, mCRP-1 with a FLAG epitope fused to is C-terminus was transiently expressed in COS cells. COS cells were transfected with the mCRP-1 cDNA using a polycationic liposome transfection reagent (Lipofectamine, GibcoBRL). Cells grown to approximately 70-80% confluence in 100 mm petri-dishes were washed in serum free DMEM media then exposed to a mixture of mCRP-1 cDNA and lipofectamine diluted in DMEM according to the manufacturers instructions. After 5 hrs incubation at 37°C with 5% CO_2 the cells were washed once with DMEM and incubated for a further 16 hrs in DMEM supplemented with 10% v/v FCS, glutamine and antibiotics (DM10). At this time the DM10 was removed and replaced with a further 8 ml/ dish of fresh DM10 and transfected cells incubated for a further 48 hrs. Supernatents containing secreted mouse CRP were recovered, centrifuged and filtered to remove cell debris, then stored at 4°C.

Protein expression

Expression of m CRP-1 from transfected COS was analysed by biosensor analysis using the M2 sensorchip (as above). mCRP-1 was purified from positive samples by M2 affinity chromatography and then purified fractions analysed by SDS-PAGE and western blot analysis. Preliminary data indicate mCRP-1 under reducing conditions is a 30 kD protein (Figure 8).

EXAMPLE 5

Expression Pattern of mCRP-1

Expression of mCRP-1 has been studied in mouse embryos. The expression pattern is similar to that of frog cerberus, with some differences, mCRP-1 is first expressed at the onset of gastrulation in midline anterior endoderm extending to the embryonic/extra-embryonic boundary. This expression is likely to be in primitive endoderm, which will eventually be displaced into the extra-embryonic region. By mid-gastrulation, expression is also seen in anterior-lateral endoderm, tissue which is likely to be the migrating wings of definitive endoderm which displaces the primitive endoderm. By the end of gastrulation, endoderm and mesoderm in the anterior half of the embryo expresses mCRP-1. Endoderm underlying the cardiac progenitors seems not to express mCRP-1. Shortly afterwards, the anterior expression fades, at first laterally, then from the midline, then is lost completely.

The very early and transient expression of mCRP-1 within anterior endoderm is very similar to that of *Xenopus* cerberus, and it is anticipated that mCRP-1 and cerberus are functionally homologous. Increasing evidence implicates midline anterior endoderm in patterning the midbrain and forebrain (5). mCRP-1 is, therefore, a good candidate for a inductive factor that either induces, or gives anterior character to, early neural tissue. As such, it may be clinically useful as an inductive, maintenance, survival, proliferative, anti-proliferative or differentiation factor in any degenerative neuropathy or in any grafting procedure involving foetal or other neural tissue grafted to correct familial or acquired deficiencies, or to repair neural tissue after trauma. Anterior lateral endoderm is also known to be the source of inducers and suppressors of cardiogenesis (4) and mCRP-1 may be one of these factors. It may be useful clinically as an inductive, maintenance, survival, proliferative, anti-proliferative or differentiation factor

for cardiomyocytes in any interventionist procedure to correct cardiomyopathy, including any grafting technique aimed at repairing infarcts (cardiomyoplasty) or valvular defects, modification (including inhibition) of familial or secondary hypertrophy, or induction of compensatory cardiomyocyte growth during aging.

Unlike *Xenopus* cerberus, mCRP-1 is also expressed in developing somites. Expression prefigures a restricted anterior compartment of the two next-to-form somites within the presomitic mesoderm, and occurs in a similar compartment of the two most recently formed somites. This expression is extremely transient and does not occur in more mature somites. mCRP-1 may be involved in establishing compartments within newly forming somites. In developing chick embryos neural crest migrates through only the anterior portion of the somite. Hence, mCRP-1 may initially give anterior character to somites, defining its interaction with neural crest. mCRP-1 may, therefore, be useful as an inductive factor in imparting particular positional qualities to engrafted tissues.

EXAMPLE 6

Biological activity of CRP

RNA encoding mCRP-1 or mCRP-2 was injected into first cleavage stage Xenopus embryos and resulted in induction of ectopic cement gland formation at later embryonic stages. As cement gland formation is often an indicator of the presence of impending appearance of neural tissue, it is likely that mCRP-1 has the capability of inducing anterior neural tissue. This conjecture is entirely consistent with the established properties of *Xenopus* cerberus. The capacity of mCRP-2 to mimic the action of *Xenopus* cerberus further suggests that this molecule has an underlying biological significance as a secreted molecule and not as transcription factor.

EXAMPLE 7

Stable Production and Characterisation of mCRP-1 Protein

A cDNA fragment containing the entire coding sequence of mCRP-1 was cloned into the expression vector pEFBOS-S-FLAG for production of C-terminal FLAG tagged protein. Transient expression following transfection of COS cells resulted in a low yield of predominantly aggregated material which, following purification by affinity and size exclusion chromatography, was found to be heterogenous at the N-terminus, most likely as a result of adverse proteolytic activity. For stable long term production, the construct and a vector incorporating a gene encoding puromycin resistance were co-transfected into CHO cells (DMEM, 10% v/v FCS, glutamine, penicillin, streptomycin, 37°C, 10% v/v CO_2) using Lipofectamine (Gibco BRL, USA) according to the manufacturer's instructions. Following selection in puromycin (25 µg/ml, Sigma, USA), resistant colonies were picked by micro-manipulation, expanded and assayed for mCRP-1-CFLAG production by binding to immobilised M2 antibody (Kodak Eastman, USA; Biosensor 2000, Pharmacia, Sweden). Several potential candidate clones were identified with clone (CL) 47 selected for further analysis. CL47 was recloned by limit dilution.

For protein production and characterisation, CL47 was expanded into Roller bottles and cultured until 3 days post confluence. An amount of 2.5 L of conditioned media was concentrated tenfold (Easy flow diafiltration apparatus, 10 kDa cut-off, Sartorius, USA). Concentrated mCRP-1 was applied to 2 ml of M2 affinity resin (Kodak Eastman) and the unbound fraction reloaded onto the column 4 times prior to extensive washing in tris-buffered saline (TBS, 500 ml). Elution (4 x 5ml) was performed using FLAG peptide (60 µg/ml, Kodak Eastman) in TBS. Fractions were monitored by SDS PAGE and Western Blot analysis for appropriate pooling of samples for further purification. Pooled M2 fractions were further purified by RP-HPLC using a Brownlee C8 column (100 x 2.1 mm I.D; Perkin Elmer, USA). Chromatography was developed using the SMART™ SYSTEM (Pharmacia, Sweden) at a flow rate of 100 µl / min. A linear gradient formed between 0.15% v/v Triflouroacetic acid (TFA) and 0.13% v/v TFA containing 60% v/v n-propanol over 60 minutes with positive fractions monitored by biosensor analysis and SDS-PAGE.

SDS PAGE and Western blot analysis of the purified protein showed the majority of mCRP-1 to have an apparent molecular weight of 38kDa under reducing conditions. Following removal of carbohydrate the protein migrated with an apparent molecular weight of 34 kDa (Panel A, Fig 9; N-glycosidase F, Boehringer Mannheim, Germany). An identical molecular weight was also apparent under non-reducing conditions suggesting that mCRP-1 is secreted by CHO cells as a monomeric protein (not shown). Size exclusion analysis (Superose 12, 300 x 3.2 mm I.D, 100 µl / min; Pharmacia) and subsequent linear regressional analysis of peak fractions compared to standards confirmed this observation (Panel B & C, Fig 9). The RP-HPLC purified mCRP-1 was subjected to N-terminal sequence analysis (Hewlett Packard, USA). The resultant 30 cycles of N terminal sequence generated a single sequence with the indicated N-terminus at D_{41} (Panel D, Figure 9). The sequence was consistent with the translated cDNA sequence.

EXAMPLE 8

Anti-BMP-like activity of mCRP-1 in Xenopus animal caps

Cerberus has been shown to have an anti-BMP-like activity when expressed in *Xenopus* animal cap assays, inducing Otx-2, a marker of anterior embryonic structures, as well as markers of neural tissue, cement gland and endoderm. To examine whether mCRP-1 shared this property, the activities of mCRP-1 and cerberus were compared in animal cap assays after injection of synthetic mRNAs at the fertilised egg stage. First, formation of cement glands in individual animal caps was scored at stage 35, since these were easily recognisable as superficial darkly pigmented patches secreting a sticky exudate. Both native mCRP-1 and CFLAG-mCRP-1 were able to induce pigmented cement glands (Figure 14) with equal frequency (Figure 14), although at only one quarter the frequency of cerberus.

Next, RT-PCR analysis was used to assess expression of various markers in pools of 25-30 animal caps derived from injected and uninjected embryos (Figure 14). Both mCRP-1 and cerberus mRNA s showed near identical activities. mCRP-1 induces the pan-neural maker *N-CAM* (9), although rather weakly, the cement gland marker *CG13* (10) and the endoderm-enriched marker Edd (11). Induced tissues apparently had anterior character, since Otx-2, a

marker of anterior embryonic structures such as forebrain, midbrain and cement gland was strongly induced, while Krox20, a hindbrain marker was not. The related homeobox genes, XNkx-2.5 and -2.3, normally expressed in cardiogenic mesoderm and anterior (pharyngeal) endoderm were strongly induced. There was no induction of the cardiac-specific differentiation marker, XMLC2a, even when caps were cultured beyond the equivalent of stage 40. Thus, the homeobox markers may be expressed in endoderm, in which case they would indicate foregut character (11). However, it is also possible that early stages of cardiogenesis are activated, without realization of the whole program.

The inductive activities of mCRP-1 and cerberus in animal caps are characteristic of a partial inhibition of BMP signalling. In animal caps, BMPs can act as both mesoderm-inducing and ventralising agents, whilst strongly inhibiting formation of neural tissue. The inventors found that BMP4 induced expression of T4 globin and XeHAND makers of ventral and lateral mesoderm, respectively (Figure 14C). To examine the relationship between cerberus and BMP4 signalling and to further compare the activities of the mouse and frog proteins, mCRP-1 and cerberus were co-expressed with BMP4 in animal caps by coinjection of equimolar amounts of their mRNAs into eggs. Both mCRP-1 and cerberus antagonised BMP4 responses, whilst BMP4 antagonised mCRP-1 and cerberus responses (Figure 14C). Thus, markers induced specifically by mCRP-1 and cerberus but not BMP4 (XNkx-2.3, XNkx-2.5, NCAM, Otx2, CG13, Edd) were reduced, as were those specifically induced by BMP4 but not mCRP-1 or cerberus (T4 globin, XeHAND). It was found that coinjected BMP4 mRNA extinguished cerberus-induced NCAM and CG13. However, other markers, XNkx-2.3, XNkx-2.5 and Edd (Xcer-induced) and T4 globin and XeHAND (BMP4-induced), were diminished but not extinguished, leaving open the possibility that cerberus and BMP4 act in a mutually antagnostic way through the same pathway. Although only semi-quantitative, the RT-PCR results indicate that BMP4 more readily overrides the inductive effects of mCRP-1 *versus* those of cerberus (Figure 14C), consistent with the mouse gene being less potent in the animal cap assay.

EXAMPLE 9

mCRP-1 expression in gastrulating embryos

Expression of mCRP-1 was examined during post-implantation development by *in situ* hybridization using a wholemount protocol and digoxygenin-labelled RNA probes, mCRP-1 transcripts were first evident at or just before the onset of gastrulation in a stripe, several cell diameters in width, along one side of the egg cylinder. This stripe extended proximally from the embryonic/extra-embryonic junction to just short of the distal tip (15A) and was on the opposite side of the egg cylinder to the primitive streak, abutting the future anterior region of epiblast fated to form anterior neurectoderm. This location was confirmed by hybridising early streak embryos with probes for both mCRP-1 and brachyary (bra) which makes prospective mesodermal cells immediately after passage through the streak (Figure 15B).

EXAMPLE 10

mCRP-1 expression in nascent somites

mCRP-1 expression faded completely from the anterior region during headfold stages. When expression was reduced to a trace at the anterior ventral midline (Figure 16A), two stripes became apparent in anterior paraxial mesoderm (Figure 16A, B). In progressively older embryos, expression was detected in 3 or 4 stripes at more and more posterior positions along the axis (Figure 16), with no other regions expressing the gene, even as late as E12.5 (Figure 16F). Histological sections of an E9.5 embryo revealed that the paraxial stripes spanned the zone of new somite formation (Figure 16E). Two stripes were restricted to the rostral half of the two most recently formed somites, while the additional stripes were positioned within proximal presomitic mesoderm. Transverse sections showed that within the cranial aspect of a single somitic stripe, expression occurred in all cells (Figure 16D).

EXAMPLE 11

mCRP-1 expression in adult tissues

mCRP-1 expression was analysed in adult tissues by RNase protection, using an mCRP-1

specific probe and a cyclophilin probe to control RNA recovery and integrity. Under conditions employed, no high level expression was detected in adult tissues including brain, skeletal muscle, salivary gland, tongue, thymus, heart, lung, stomach, spleen, liver, pancreas, intestine, kidney, bladder, uterus, testes and ovary.

EXAMPLE 12

mCRP-1 expression in Otx-2$^{-/-}$ embryos

Anterior primitive endoderm is known to be essential for patterning the anterior epiblast and neural plate as is anterior embryonic mesendoderm at later stages. Since mCRP-1 is expressed in both of these zones, it could participate in the patterning function, either as an inducer, or an inhibitor. Another gene that may be involved is *Otx-2*, which encodes a homeodomain protein related to *Drosophila* orthodenticle and empty spiracles, both involved in head development in the fly. *Otx-2* is initially expressed throughout the epiblast of the gastrula, before becoming restricted to anterior tissues including chordamesoderm, forebrain and midbrain. Targeted *Otx-2*$^{-/-}$ mutants show severe abnormalities at gastrulation, including defective formation of chordamesoderm and loss of all brain compartments anterior to rhombomere 3. A possible role for *Otx-2* in the patterning function of primitive endoderm was suggested by examination of null embryos carrying a LacZ reporter gene, which showed that while *Otx-2* could be expressed in primitive endoderm in the mutant context, it could not be induced and/or maintained in anterior neural plate.

mCRP-1 expression was examined in *Otx-2*$^{-/-}$ embryos by wholemount *in situ* hybridisation (Figure 17) and found that it was still expressed in an anterior region, although to varying extents. In more severely affected embryos, expression was restricted to the distal tip and was absent at the embryonic/extra-embryonic junction, marker by a prominent constriction. As noted above, mCRP-1 would normally extend up to this point. In the less affected *Otx-2*$^{-/-}$ mutant examined (Figure 17), mCRP-1 expression tended more towards the norla pattern, although, since anterior patterning is still disturbed in these embryos, it is difficult to say to what extent.

EXAMPLE 13

Stable Production and Characterisation of mCRP-2 Protein

A cDNA fragment containing the entire coding sequence of mCRP-2 was cloned into the expression vector pEFBOS-S-FLAG for production of C-terminal FLAG tagged protein. Transient expression following transfection of COS cells resulted in good yields of predominantly dimeric mCRP-2. For stable, long term production the construct and a vector incorporating a gene encoding puromycin resistance were co-transfected into rat SR-3Y1 cells (DMEM, 10% v/v FCS, glutamine, penicillin, streptomycin, 37°C, 10% v/v CO_2) using Lipofectamine (Gibco BRL, USA) according to the manufacturers instructions. Following selection in puromycin (25 µg/ml, Sigma, USA), resistant colonies were picked by micro-manipulation, expanded, and assayed for mCRP-2-CFLAG production by binding to immobilised M2 antibody (Kodak Eastman, USA; Biosensor 2000, Pharmacia, Sweden). Several potential candidate clones were identified with clone (CL) 23 selected for further analysis. CL23 was recloned following culture and micro-manipulation from soft agar.

For protein production and characterisation CL23 was expanded into Roller bottles and cultured until 3 days post confluence. An amount of 2.5 L of conditioned media was concentrated tenfold (Easy flow diafiltration apparatus, 10 kDa cut-off, Sartorius, USA). The sample was also buffer exchanged after concentration into 20mM Tris, 0.15% w/v NaCl, 0.1% v/v Tween 20 (TBS). Concentrated mCRP-2 was applied to 2 ml of M2 affinity resin (Kodak Eastman) and the unbound fraction reloaded onto the column 4 times prior to extensive washing in Tris-buffered saline (TBS, 500 ml). Elution (4 x 5ml) was performed using FLAG peptide (60 µg/ml, Kodak Eastman) in TBS. Fractions were monitored by SDS PAGE and Western Blot analysis for appropriate pooling of samples for further purification.

Pooled M2 fractions were applied to G25 Sepharose resin (Pharmacia) packed into an XK column (400 x 26mm I.D.) to remove free peptide from the previous step. The column was operated at 4ml/min. using an FPLC with online UV_{280} absorbance and conductivity detection.

The column buffer was 20mM Tris, 0.1M NaCl, 0.02%v/v Tween 20, 0.05% w/v azide pH 7.5. Fractions were collected at 2.5 min. intervals and monitored on the biosensor. Peak fractions were diluted 1:1 with 20mM Tris, 0.02%v/v Tween 20, 0.05% w/v azide pH 7.5 (Buffer A) and applied to a previously equilibrated MonoQ 5/5 column (Pharmacia, 50 x 5 mm I.D.) via a superloop. After sample loading was complete, the column was re-equilibrated into Buffer A and a gradient developed over 50 min from Buffer A to 1.0M NaCl in Buffer A. Fractions containing mCRP-2 were monitored by biosensor analysis and SDS PAGE. Typically 400µg of mCRP-2 was recovered in fractions 18/19 which eluted at 0.35-0.4 M NaCl and was essentially pure (>95%).

Protein estimations were determined by SDS PAGE analysis (comparison to known standard concentrations), Western blot, and also by comparison of 10µg Bovine serum albumin under identical ion exchange chromatographic conditions. Reducing conditions on SDS PAGE (0.05 % v/v β mercaptoethanol) resulted in a single band residing at 27 kD (Panels A and B, Fig 10). Following removal of carbohydrate the protein was reduced to an apparent molecular weight of 24 kDa (Panel A, Fig 10; N-glycosidase F, Boehringer Mannheim, Germany). Under non-reducing conditions a single band residing at approximately 55kD was apparent (Panel B, Figure 10) suggesting that under native conditions mCRP-2 exists as a homodimer. Size exclusion analysis of 1 µg of high purity mCRP-2 (Superose 12, 300 x 3.2 mm I.D, 100 µl / min; Pharmacia) and subsequent linear regressional analysis of peak fractions compared to standards confirmed this observation (Panel D and C, Fig 10). High purity DAN (5 µg) was subjected to N-terminal sequence analysis (Hewlett Packard, USA). The resultant 30 cycles of N terminal sequence generated a single sequence with the indicated N-terminus at Ala_{17} (Panel E, Figure 10). The sequence was consistent with the translated cDNA sequence and identified a N linked glycosylation site at N_{38}.

EXAMPLE 14

Isolation of a genomic DNA encoding human CRP-1

A human genomic DNA clone encoding a human CRP (hCRP-1) was isolated by screening a human genomic library (lambda DASH II, Stratagene) with cDNA encoding mCRP-1. Library screens were performed using hybridisation conditions described by Sambrook *et al* (6). Hybridisation and washing conditions were performed at relatively high stringency at 65°C and 65°C, 2 xSSC/0.1% w/v SDS, respectively. The probe was generated by random priming of the cDNA encoding mCRP-1.

The genomic clone, hCRP-1 was sequenced and found to contain a full open reading frame, encompassing two exons and one intron and parts of the 5′UTR and 3′UTR. The sequence data obtained were used to construct a continuous 3,150 base pair sequence from which the putative full length protein could be translated. The complete genomic nucleotide sequence and corresponding amino acid sequence to exons 1 and 2 is shown in Figure 11. The putative coding sequence without intron and corresponding protein translation is shown in Figure 12. Although initial Southern blot analysis using hCRP-1 as a probe on mouse genomic DNA indicated that this gene was not the mCRP-1 equivalent, more recent analysis now shows that hCRP1 is the mCRP-1 equivalent.

The translated sequence shows sequence similarities with mCRP-1 and *Xenopus laevis* cerberus. A protein comparison of hCRP-1, mCRP-1 and *Xenopus laevis* cerberus is shown in Figure 13.

Southern blot/PCR analysis on a panel of CHO human hybrids indicated that the hCRP-1 gene maps to 9q23 which is syntenic with the central region of mouse chromosome 4 that harbours the mCRP-1 gene.

EXAMPLE 15

Chromosome location of mCRP-1

The mouse chromosomal location of mCRP-1 was determined by interspecific backcross analysis of a panel of progeny derived from matings of [(C57BL/6J/ x *Mus spretus)* F1 x C57BL/6J] mice. This interspecific backcross panel has been typed for over 2400 loci that are well distributed among all autosomes as well as the X chromosome (12). DNA isolation, restriction enzyme digestion, agarose gel electrophoresis, Southern transfer to Hybond-N+ and hybridisations were performed essentially as described (13). The probe for mCRP-1 encompassed nucleotides 288-640 and was radiolabelled with [^{32}P]-CTP using a random priming kit (Stratagene). Washing of Southern filters was performed in 1 x SSCP, 0.1% w/v SDS at 65°C. A major fragment of 6-kb was detected in *Sac*I-digested C57BL/6J DNA and a major fragment of 3-kb was detected in *Sac*I-digested *M. spretus* DNA. The presence or absence of the 3-kb *SacIM. spretus* fragment was followed in backcross mice. A description of probes and RFLPs for loci linked to mCRP-1 has been (14, 15) previously reported. Recombination distances were calculated using the Map Manager program (version 2.6.5). Gene order was determined by minimizing the number of recombination events required to explain the allele distribution patterns.

The results indicate that mCRP-1 is located in the central region of mouse chromosome 4, linked to *Tyrp1, Ifna, Jun* and *Pgm2* (Figure 18). Ninety-two mice were analysed with all of these markers, although for some marker pairs up to 175 mice were typed. To calculate recombination frequencies, each locus was analysed in pairwise combinations using the additional data. The ratios of the total number of mice exhibiting recombinant chromosomes to the total number of mice analysed for each pair or loci, and most the likely gene order are: centrometer - *Tyrp1* - 1/163 -*mCer*-1 2/116 - *Ifna* - 0/122 - *Jun* -5/174 - *Pgm2*.

Recombination frequencies (expressed as genetic distances in centiMorgans (cM)±standard error) are *Tyrp1* - 0.6±0.6 - *mCer-1* - 2.6±1.5 - [*Ifna, Jun*] - 2.9±1.3. No recombinants were detected between *Ifna* and *Jun* in 122 animals types, suggesting that the two loci are within 2. 5 cM of each other (upper 95% confidence limit).

The inventors have compared the interspecific map of chromosome 4 with a composite mouse linkage map that reports the location of many uncloned mutations (provided by the Mouse Genome Database, a computerised databased maintained at the Jackson Laboratory, Bar Harbor, ME, USA). A number of mouse mutations map with 5 cM of the chromosomal position determined for mCRP-1: pintail (pt) (16), polysyndactyly (ps) (17, 18), meander tail (mea) (19) and head blebs (heb) (20). All of these mutations result in phenotypes which include defects in head or tail development and could conceivably be associated with mutation of mCRP-1 or perturbation of its expression. To test this possibility, genomic DNA derived from heterozygous or homozygous individuals of each mutant strain was analysed by Southern blotting using an mCRP-1 probe that detected both coding exons (Figure 18). 10μg of each DNA was digested separately with *Eco*RI and *Bam*HI, resolved by agarose gel electrophoresis and Southern blotted. After hybridisation, filters were washed in 0.1 x SSC, 0.1% w/v SDS at 65°C. For *Eco*RI digests, major fragments of 4.5 and 5.5 kb were detected in wildtype and all mutant strains except pt, in which a polymorphism resulted in a 5.5-kb doublet (Figure 18). Since the pt sample analysed was from a heterozygote animal, complete loss of the 4.5-kb fragment indicated that the observed polymorphism was not specific to the pt allele. For *Bam*HI digests, a single 4.3-kb fragment was detected. The Southern analysis showed that the mCRP-1 locus is not grossly rearranged in the mutant strains. If total deletion of the mCRP-1 locus had occurred in stains for which heterozygous DNA was analysed (pt and ps), the Southern banding pattern produced would be the same, although the signal would be half the intensity. This was shown not to be the case by controlling for DNA loading with a probe corresponding to the homeobox gene *Nkx*2-5 (21), located on chromosome 17 (22). To examine the possibility that small deletions or point mutations in the mCRP-1 coding region were responsible for any mutant phenotype, the inventors amplified the mCRP-1 coding region from each strain by polymerase chain reaction and determined its DNA sequence directly. The results showed no candidate mutations. Within homozygote samples (mea and heb), no base changes that led to amino acid changes were detected. For heterozygote samples (ps and pt), a candidate mutation would manifest as an ambiguous base. No such ambiguities were detected, although a few polymorphisms were present. The inventors conclude that gross rearrangement or coding region mutations in mCRP-1 do not account for any of the mutant phenotypes.

EXAMPLE 16

mCRP-1 is a secreted N-terminally processed glycoprotein

Methods

pEFBOS I FLAG vector carrying the mCRP-1 insert was co-transfected with a puromycin resistance plasmid into CHO cells using Lipofectamine (Gibco BRL) according to manufacturer's instructions. Following puromycin selection (Sigma; 25mg/ml), resistant clones were picked and expanded, and culture supernatant was assayed fro CFLAG-mCRP-1 by analysis of binding to immobilised M2 (anti-FLAG) antibody (Kodak Eastman) on a Biosensor 2000 (Pharmacia). CFLAG-mCRP-1 was purified from conditioned medium of clone CL47 using M2 affinity resin (Kodak Eastman), eluding with FLAG peptide (60 μg/ml; Kodak Eastman) in Tris-buffered saline. Fractions were monitored by Western blotting probed with M2 antibody. Pooled fractions were further purified by reversed phase-HPLC using a Brownlee C8 column (100 mm x 2.1 mm I.D.; Perkin Elmer, USA) on a SMART system (Pharmacia) using a flow rate of 100 μl/min and a linear gradient between 0.15% (v/v) trifluoroacetic acid (TFA) and 0.13% (v/v) TFA containing 60% (v/v) n-propanol. Fractions were monitored on the Biosensor and by SDS-PAGE. For N-glycosylation analysis, purified protein (1μg) was treated with N-glycosidase F (Boehringer), as per manufacturer's instructions; before Western blotting. Purified material, subjected to automated unambiguous sequence over 30 cycles concordant with the predicted open reading frame of mCRP-1.

293T fibroblast cells were transfected (Lipofectamine) with an pEFBOS construct encoding a C' terminal flag tagged mCRP-1. Three days post transfection supernatant was harvested and expression assessed by biosensor analysis. mCRP-1-C′ was affinity purified M2 (anti-falg antibody) resin (1ml). Flag peptide was used to elute mCRP-1 from the column at a concentration of 100μg/ml (5ml). Purified fractions were monitored by Biosensor and by SDS-PAGE. Preliminary Western analysis indicated a 38kDa polypeptide under reducing conditions and a 74kDa polypeptide under non-reducing conditions. N-terminal amino acid sequence analysis of mCRP-1 derived from 293T cells was done and the sequence was the same as that derived from mCRP-1 expressed from CHO cells.

Results

A PCR-generated cDNA fragment spanning the mCRP-1 coding region was cloned into the expression vector pEFBOS I FLAG for production of recombinant protein carrying a C-terminal FLAG epitope. CFLAG-mCRP-1 could be recovered from culture supernatans of both Chinese hamster ovary (CHO) cells, after stable integration of vector, as well as from 293T cells after transient transfection (Fig. 20A, B). CFLAG-mCRP-1 purified from CHO cells using immobilised M2 (anti-FLAG) antibody and reversed phase-HPLC migrated principally at 38kD on reducing and denaturing gels. It size was reduced to 32kD after treatment with N-glycosidase F (Fig. 20A), demonstrating that mCRP-1, as expected of a secreted protein, is N-glycosylated. To assess whether CFLAG-mCRP-1 was processed, N-terminal amino acid sequence analysis was performed on purified material. The sequence unambiguously showed cleavage after the basis peptide RGRR, at a position 40 amino acids into the native protein. However, examination of CHO cell preparations on denaturing and non-reducing gels, and by gel filtration, indicated that most unaggregated protein migrated as monomer, with little dimer detected. Since all known cystine knot cytokines are secreted as homo- or heterodimers (9), CHO cell-derived material may be abnormally processed or secreted. The inventors examined, CFLAG-mCRP-1 secreted from 293T cells, and in this case found dimers, with no monomer detected (Fig. 20B). These data demonstrate that mCRP-1 is a secreted, N-terminally processed glycoprotein that can form predominantly disulphide bonded dimers when secreted from certain cell types.

EXAMPLE 17

A gene encoding hCRP-1 was isolated from a human genomic library (see Example 14). PCR was used to generate probes specific for putative hCRP-1 exons 1 and 2 which were used to screen a number of cDNA libraries constructed from various fetal and adult tissues. No cDNA clones encoding hCRP-1 were isolated.

To express hCRP-1 protein, an artificial hCRP-1 cDNA was constructed using standard techniques such as splice overlap extension PCR (SOE-PCR). Two separate PCR's were

performed to generate exon 1 and exon 2 specific hCRP-1 cDNA, the PCR fragments purified and used in a third PCR to generate a full length cDNA from which the intron has been excised. Attempts were made to clone this cDNA into the expression vector pEFBOS I FLAG for production of a recombinant protein carrying a C-terminal FLAG epitope. This construct proved to be unstable so other epitope tags were employed. A stable using the pEFBOS vector with an I-SPY epitope tag (peptide seq: LNQYPALTE; AMRAD Biotech) at the C-terminus of hCRP-1 was generated (C'-I-SPY-hCRP-1).

For transient expression of hCRP-1, 293T human fibroblast cells were transfected (Lipofectamine, Gibco BRL) with the pEFBOS I-SPY-hCRP-1 construct according to the manufacturers instructions. Three days post transfection supernatant was harvested and expression assessed by biosensor analysis where I-SPY antibody (D11) was immobilised to the sensorchip at 100 ng/ml. C'-I-SPY-hCRP-1 was affinity purified from the supernatant using an I-SPY antibody coupled resin (AMRAD Biotech). Purified material was analysed by Western blot analysis using the I-SPY antibody conjugated with HRPO and subsequent ECL detection. Results presented in Fig 21 demonstrate that under reducing conditions the major species of protein was a monomer migrating with an apparent molecular weight of approximately 40 kDa. Other minor species present under reducing conditions are non-reduced dimers, as well as a small amount of aggregated material. Under non-reducing conditions the monomeric form was replaced by a dimer of approximately 80 kDa and a large amount of aggregated material. These results indicate that like mCRP-1 and other cystine knot cytokines, hCRP-1 exists primarily as a dimeric molecule.

To further confirm the relationship between hCRP-1 and mCRP-1, experiments were undertaken to determine if the human and mouse molecules would from heterodimers when simultaneously expressed in the same cells. To this end 293T cells were cotransfected with vectors encoding C'-FLAG-mCRP-1 and C'-I-SPY-hCRP-1 respectively. In a number of experiments the I-SPY specific antibody was shown to immunoprecipitate a protein of the appropriate molecular weight that contained a FLAG epitope tag as assessed in Western blot analysis. These results indicated that hCRP-1 and mCRP-1 monomers will, under appropriate conditions, interact to form disulphide linked heterodimers.

EXAMPLE 18

EXPRESSION PATTERNS OF mCRP-2

In the mouse embryo, mCRP-2 is expressed from the beginning of organogenesis (embryonic day 8.5) within the segmenting paraxial mesoderm. mCRP-2 expression prefigures the next-to-form somite within unsegmented mesoderm and in the most recently formed somite. Thus, both mCRP-2 and mCRP-1 are expressed during somite formation, and may be involved in the same type of process. In contrast to mCRP-1, however, mCRP-2 expression is strongest in the posterior region of somites. mCRP-2 and mCRP-1 may therefore act in compartmentalisation of the posterior and anterior regions of the somite, respectively. After dropping considerably in newly formed somites, mCRP-2 expression increases as somites mature. Somites differentiate into three compartments along the dorso-ventral axis, fated to give rise to the dermis dorsally (dermatome), the back and limb muscles (myotome) and the vertebrae ventrally (sclerotome). mCRP-2 expression in somites is strongest medially and dorsally. mCRP-2 is also expressed in the limb musculature and in craniofacial epithelium and mesenchyme.

These expression domains suggest a number of roles for mCRP-2 in early embryos. First, in somite patterning, perhaps performing an analogous role to mCRP-1. Secondly, in muscles of the body and limbs. The function of mCRP-2 may be tissues of the somite, the dermatome (forming skin) and schlerotome (forming bone). Thirdly, in craniofacial development. mCRP-2 may be useful clinically as an inductive, maintenance, survival, proliferative, anti-proliferative or differentiation factor in pathologies related to muscle, bone and skin. Furthermore, its proposed role as a tumour suppressor suggests a function as an anti-proliferative factor in a broad range of cancers, including breast cancer, lymphoma and leukaemia, melanoma, colorectal cancer, pancreatic cancer, lung cancer, stomach cancer and nueroblastoma.

mCRP-2 is expressed in all adult tissues examined with the exception of the liver. Strongest expression levels were observed in bladder, uterus and lungs.

EXAMPLE 19

mCRP-1 AND mCRP-2 FORM HETERODIMERS

Results indicated that mCRP-1 and mCRP-2 are capable of either associating, or existing in a heterodimeric complex, with each other. This possibility is raised by analysis of supernatants recovered from cells which had been transfected with expression vectors for mCRP-1 and mCRP-2. Briefly, DNA expression vectors designed to produce c-terminally flag tagged mCRP-1 and c-terminally myc tagged mCRP-2 were transfected, either individually or together, into 293T cells using the Lipofectamine transfection protocol described previously. 72 hours later, the supernatants from the various transfections were harvested, clarified by low speed centrifugation and then treated in two different ways. First, 5 μl samples of each supernatant was analysed by SDS PAGE and Western blotting to assess the level of mCRP-1-flag and mCRP-2-myc proteins. Second, 1 ml of the supernatant was incubated with 10 μl of M2 anti-flag antibody affinity resin to bind proteins containing the flag epitope. The resin was washed a number of times prior to elution of bound proteins with 20 μl of 5 μg/ml M2 peptide. The eluate from this procedure was also subjected to analysis by SDS PAGE and western blotting. In all the procedures the western blots were probed with either the anti-flag M2 antibody or the 9E10 anti-myc antibody. Following binding of the HRP-conjugated secondary antibody, the proteins were then detected using ECL.

This analysis showed that the amount of mCRP-2-myc recovered from the flag affinity purified mCRP-1/mCRP-2 supernatant was significantly more than that recovered from a parallel samples from which the mCRP-1-flag protein was absent. This result suggests that mCRP-2-myc has copurified with mCRP-1-flag raising the possibility that the two proteins either associate of dimerize. Given both proteins share a degree of structural similarity it would not be unusual if the proteins could form heterodimers. Heterodimeric molecules are commonly observed with the TBF-β class of cystine knot proteins.

The existence of heterodimers between mCRP-1 and -2 raises the following possibilities. First, the activity of the heterodimer may be different from homodimers of either mCRP-1 or -2. A heterodimer may be able to bind to distinct receptors, if they exist, and elicit novel responses

in the cells that possess them. A heterodimer may possess distinct biochemical properties, in terms of its stability and or solubility. It is also possible that the biological potency of a heterodimer may exceed that of molecules being homodimeric examples of its constituents.

Those skilled in the art will appreciate that the invention described herein is susceptible to variations and modifications other than those specifically described. It is to be understood that the invention includes all such variations and modifications. The invention also includes all of the steps, features, compositions and compounds referred to or indicated in this specification, individually or collectively, and any and all combinations of any two or more of said steps or features.

BIBLIOGRAPHY

1. Bouwmeester *et al. Nature* 382: 595-601,1996.
2. Ozaki and Sakiyama *Proc. Natl. Acad-Sci. USA 90:* 2593-2597, 1993.
3. Enomoto *et al Oncogene 9:* 2785-2791, 1994.
4. Schultheiss, *et al. Genes Dev. 11:* 451-462, 1997.
5. Thomas and Beddington. *Current Biology 6:* 1487-1496, 1996.
6. Sambrook *et al. Cloning: A Laboratory Manual.* Cold Spring Harbour, NY, USA; Second Edition; 1989.
7. McDonald and Hendrickson *Cell 73:*421-424, 1993.
8. Isaacs *Current Opinion in Structural Biology 5:*391-395, 1995.
9. Kinter and Melton *Development 99:* 311-325, 1987.
10. Jamrich and Sato *Development 105:* 779-786, 1989.
11. Sasai *et al EMBO J 15:* 4547-4555, 1996.
12. Copeland and Jenkins *Trends Genet 7:* 113-118, 1991.
13. Jenkins *et al J. Virol 43:* 26-36, 1982.
14. Smith *et al Cell 73:* 1349-1360, 1993.
15. Ceci *et al Genomics 5:* 699-709, 1989.
16. Hollander and Strong *J. Hered 42:* 179-182, 1951.
17. Batchelor *et al Mutation Research 3:* 218-229, 1968.
18. Johnson *J Embryol. Exp. Morphol. 21:* 285-294, 1969.
19. Hollander and Waggie *J. Hered 68:* 403-406, 1977.
20. Varnum and Fox *J. Hered 72:* 293, 1981.
21. Lints *et al Development 119:* 419-431, 1993.
22. Himmelbauer *et al Mammalian Genome 5:* 814-816, 1995.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: AMRAD OPERATIONS PTY LTD
OTHER THAN US: HILTON Douglas J, STANLEY, Edouard G, HARVEY Richard P, BIBEN Christine, FABRI Louis, LAH Maria and NASH Andrew D

(ii) TITLE OF INVENTION: NOVEL MOLECULES AND USES THEREOF

(iii) NUMBER OF SEQUENCES: 23

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: DAVIES COLLISON CAVE
(B) STREET: 1 LITTLE COLLINS STREET
(C) CITY: MELBOURNE
(D) STATE: VICTORIA
(E) COUNTRY: AUSTRALIA
(F) ZIP: 3000

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: PCT INTERNATIONAL (PCT)
(B) FILING DATE: 11-FEB-1998
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: PO8963
(B) FILING DATE: 3-SEP-1997
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: PO5067
(B) FILING DATE: 11-FEB-1997
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: PO6420
(B) FILING DATE: 24-APR-1997
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: PP0961
(B) FILING DATE: 16-DEC-1997
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: HUGHES, DR E JOHN L
(C) REFERENCE/DOCKET NUMBER: EJH/AF

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: +61 3 9254 2777
(B) TELEFAX: +61 3 9254 2770
(C) TELEX: AA 31787

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Cys Xaa Tyr Xaa Pro Phe Xaa Gln Xaa Ile Xaa His Glu Xaa Cys Xaa Xaa Xaa Val
5 10 15
Xaa Gln Asn Asn Leu Cys Phe Gly Lys Cys Xaa Ser Xaa Xaa Xaa Pro Xaa_{n1} Cys Ser
20 25 30 35
His Cys Xaa Pro
40

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Cys Arg Tyr Val Pro Phe Asn Gln Tyr Ile Ala His Glu Asp Cys Gln Lys Val Val
5 10 15
Val Gln Asn Asn Leu Cys Phe Gly Lys Cys
20 25

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 995 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 54..869

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
TTTTAGGCCC GTCCATCTGT GAATCTAACC TCAGTCTCTG GGAATCAGGA AGC ATG      56
                                                            Met
                                                             1

CAT CTC CTC TTA GTT CAG CTG CTT GTT CTC TTG CCT CTG GGG AAG GCA     104
His Leu Leu Leu Val Gln Leu Leu Val Leu Leu Pro Leu Gly Lys Ala
                5                  10                  15

GAC CTA TGT GTG GAT GGC TGC CAG AGT CAG GGC TCT TTA TCC TTT CCT     152
Asp Leu Cys Val Asp Gly Cys Gln Ser Gln Gly Ser Leu Ser Phe Pro
            20                  25                  30

CTC CTA GAA AGG GGT CGC AGA GAT CTC CAC GTG GCC AAC CAC GAG GAG     200
Leu Leu Glu Arg Gly Arg Arg Asp Leu His Val Ala Asn His Glu Glu
        35                  40                  45

GCA GAA GAC AAG CCG GAT CTG TTT GTG GCC ATG CCA CAC CTC ATG GGC     248
Ala Glu Asp Lys Pro Asp Leu Phe Val Ala Met Pro His Leu Met Gly
    50                  55                  60                  65

ACC AGC CTG GCT GGG GAA GGT CAG AGG CAG AGA GGG AAG ATG CTG TCC     296
Thr Ser Leu Ala Gly Glu Gly Gln Arg Gln Arg Gly Lys Met Leu Ser
                70                  75                  80

AGG CTT GGA AGA TTC TGG AAG AAA CCT GAG ACC GAA TTT TAC CCC CCA     344
Arg Leu Gly Arg Phe Trp Lys Lys Pro Glu Thr Glu Phe Tyr Pro Pro
            85                  90                  95
```

```
AGG GAT GTG GAA AGC GAT CAT GTC TCA TCG GGG ATG CAG GCC GTG ACT      392
Arg Asp Val Glu Ser Asp His Val Ser Ser Gly Met Gln Ala Val Thr
        100                 105                 110

CAG CCA GCA GAT GGG AGG AAA GTG GAG AGA TCA CCT CTA CAG GAG GAA      440
Gln Pro Ala Asp Gly Arg Lys Val Glu Arg Ser Pro Leu Gln Glu Glu
    115                 120                 125

GCC AAG AGG TTC TGG CAT CGG TTC ATG TTC AGA AAG GGC CCG GCG TTC      488
Ala Lys Arg Phe Trp His Arg Phe Met Phe Arg Lys Gly Pro Ala Phe
130                 135                 140                 145

CAG GGA GTC ATC CTG CCC ATC AAA AGC CAC GAA GTA CAC TGG GAG ACC      536
Gln Gly Val Ile Leu Pro Ile Lys Ser His Glu Val His Trp Glu Thr
                150                 155                 160

TGC AGG ACT GTG CCC TTC AAC CAG ACC ATT GCC CAT GAA GAC TGT CAA      584
Cys Arg Thr Val Pro Phe Asn Gln Thr Ile Ala His Glu Asp Cys Gln
            165                 170                 175

AAA GTC GTT GTC CAG AAC AAC CTT TGC TTT GGC AAA TGC AGT TCC ATT      632
Lys Val Val Val Gln Asn Asn Leu Cys Phe Gly Lys Cys Ser Ser Ile
        180                 185                 190

CGT TTT CCC GGA GAA GGG GCA GAT GCC CAC AGC TTC TGC TCC CAC TGC      680
Arg Phe Pro Gly Glu Gly Ala Asp Ala His Ser Phe Cys Ser His Cys
    195                 200                 205

TCG CCC ACC AAA TTC ACC ACC GTG CAC TTG ATG CTG AAC TGC ACC AGC      728
Ser Pro Thr Lys Phe Thr Thr Val His Leu Met Leu Asn Cys Thr Ser
210                 215                 220                 225

CCA ACC CCC GTG GTC AAG ATG GTG ATG CAA GTA GAA GAG TGT CAG TGC      776
Pro Thr Pro Val Val Lys Met Val Met Gln Val Glu Glu Cys Gln Cys
                230                 235                 240

ATG GTG AAG ACG GAA CGT GGA GAG GAG CGC CTC CTA CTG GCT GGT TCC      824
Met Val Lys Thr Glu Arg Gly Glu Glu Arg Leu Leu Leu Ala Gly Ser
            245                 250                 255

CAG GGT TCC TTC ATC CCT GGA CTT CCA GCT TCA AAA ACA AAC CCA           869
Gln Gly Ser Phe Ile Pro Gly Leu Pro Ala Ser Lys Thr Asn Pro
        260                 265                 270

TGAATTACCT CAACAGAAAG CAAAACCTCA ACAGAATAGG TGAGGTTATT CAATCTGGAA    929
```

ATGTTATGTG AGTTTATATA AAGATCAGTG GAAAATAAAA AAAAAAAAAA AAAAAAAAAA 989

AAAAAA 995

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 272 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met His Leu Leu Leu Val Gln Leu Leu Val Leu Leu Pro Leu Gly Lys
1 5 10 15

Ala Asp Leu Cys Val Asp Gly Cys Gln Ser Gln Gly Ser Leu Ser Phe
20 25 30

Pro Leu Leu Glu Arg Gly Arg Arg Asp Leu His Val Ala Asn His Glu
35 40 45

Glu Ala Glu Asp Lys Pro Asp Leu Phe Val Ala Met Pro His Leu Met
50 55 60

Gly Thr Ser Leu Ala Gly Glu Gly Gln Arg Gln Arg Gly Lys Met Leu
65 70 75 80

Ser Arg Leu Gly Arg Phe Trp Lys Lys Pro Glu Thr Glu Phe Tyr Pro
85 90 95

Pro Arg Asp Val Glu Ser Asp His Val Ser Ser Gly Met Gln Ala Val
100 105 110

Thr Gln Pro Ala Asp Gly Arg Lys Val Glu Arg Ser Pro Leu Gln Glu
115 120 125

Glu Ala Lys Arg Phe Trp His Arg Phe Met Phe Arg Lys Gly Pro Ala
130 135 140

Phe Gln Gly Val Ile Leu Pro Ile Lys Ser His Glu Val His Trp Glu
145 150 155 160

```
Thr Cys Arg Thr Val Pro Phe Asn Gln Thr Ile Ala His Glu Asp Cys
                165                 170                 175

Gln Lys Val Val Val Gln Asn Asn Leu Cys Phe Gly Lys Cys Ser Ser
            180                 185                 190

Ile Arg Phe Pro Gly Glu Gly Ala Asp Ala His Ser Phe Cys Ser His
        195                 200                 205

Cys Ser Pro Thr Lys Phe Thr Thr Val His Leu Met Leu Asn Cys Thr
    210                 215                 220

Ser Pro Thr Pro Val Val Lys Met Val Met Gln Val Glu Glu Cys Gln
225                 230                 235                 240

Cys Met Val Lys Thr Glu Arg Gly Glu Glu Arg Leu Leu Leu Ala Gly
                245                 250                 255

Ser Gln Gly Ser Phe Ile Pro Gly Leu Pro Ala Ser Lys Thr Asn Pro
            260                 265                 270
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 178 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
Met Leu Trp Val Leu Val Gly Ala Val Leu Pro Val Met Leu Leu Ala
1               5                   10                  15

Ala Pro Pro Pro Ile Asn Lys Leu Ala Leu Phe Pro Asp Lys Ser Ala
            20                  25                  30

Trp Cys Glu Ala Lys Asn Ile Thr Gln Ile Val Gly His Ser Gly Cys
        35                  40                  45
```

Glu Ala Lys Ser Ile Gln Asn Arg Ala Cys Leu Gly Gln Cys Phe Ser
50 55 60

Tyr Ser Val Pro Asn Thr Phe Pro Gln Ser Thr Glu Ser Leu Val His
65 70 75 80

Cys Asp Ser Cys Met Pro Ala Gln Ser Met Trp Glu Ile Val Thr Leu
85 90 95

Glu Cys Pro Asp His Glu Glu Val Pro Arg Val Asp Lys Leu Val Glu
100 105 110

Lys Ile Val His Cys Ser Cys Gln Ala Cys Gly Lys Glu Pro Ser His
115 120 125

Glu Gly Leu Asn Val Tyr Val Gln Gly Glu Asp Ser Pro Gly Ser Gln
130 135 140

Pro Gly Pro His Ser His Ala His Pro His Pro Gly Gly Gln Thr Pro
145 150 155 160

Glu Pro Glu Glu Pro Pro Gly Ala Pro Gln Val Glu Glu Glu Gly Ala
165 170 175

Glu Asp

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 178 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Leu Trp Val Leu Val Gly Ala Val Leu Pro Val Met Leu Leu Ala
1 5 10 15

Ala Pro Pro Pro Ile Asn Lys Leu Ala Leu Phe Pro Asp Lys Ser Ala
20 25 30

Trp Cys Glu Ala Lys Asn Ile Thr Gln Ile Val Gly His Ser Gly Cys
35 40 45

Glu Ala Lys Ser Ile Gln Asn Arg Ala Cys Leu Gly Gln Cys Phe Ser
50 55 60

Tyr Ser Val Pro Asn Thr Phe Pro Gln Ser Thr Glu Ser Leu Val His
65 70 75 80

Cys Asp Ser Cys Met Pro Ala Gln Ser Met Trp Glu Ile Val Thr Leu
85 90 95

Glu Cys Pro Asp His Glu Glu Val Pro Arg Val Asp Lys Leu Val Glu
100 105 110

Lys Ile Val His Cys Ser Cys Gln Ala Cys Gly Lys Glu Pro Ser His
115 120 125

Glu Gly Leu Asn Val Tyr Val Gln Gly Glu Asp Ser Pro Gly Ser Gln
130 135 140

Pro Gly Pro His Ser His Ala His Pro His Pro Gly Gly Gln Thr Pro
145 150 155 160

Glu Pro Glu Glu Pro Pro Gly Ala Pro Gln Val Glu Glu Glu Gly Ala
165 170 175

Glu Asp

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 186 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met Leu Arg Val Leu Val Gly Ala Val Leu Pro Ala Met Leu Leu Ala

1 5 10 15

Ala Pro Pro Pro Ile Asn Lys Leu Ala Leu Phe Pro Asp Lys Ser Ala
20 25 30

Trp Cys Glu Ala Lys Asn Ile Thr Gln Ile Val Gly His Phe Thr Ser
35 40 45

Gly Cys Glu Ala Lys Ser Ile Gln Asn Arg Ala Cys Leu Gly Gln Cys
50 55 60

Phe Ser Tyr Ser Val Pro Asn Thr Phe Pro Gln Ser Thr Glu Ser Leu
65 70 75 80

Val His Cys Asp Ser Cys Met Pro Ala Gln Ser Met Trp Glu Ile Val
85 90 95

Thr Leu Glu Cys Pro Gly His Glu Glu Val Phe Thr Pro Arg Val Asp
100 105 110

Lys Leu Val Glu Lys Ile Leu His Cys Ser Cys Gln Ala Cys Gly Lys
115 120 125

Glu Pro Ser His Glu Gly Leu Ser Val Tyr Val Gln Gly Glu Asp Gly
130 135 140

Pro Gly Ser Gln Pro Gly Thr His Pro His Pro His Pro His Pro His
145 150 155 160

Pro Gly Gly Gln Thr Pro Glu Phe Thr Pro Glu Asp Pro Pro Gly Ala
165 170 175

Pro His Thr Glu Glu Glu Gly Ala Glu Asp
180 185

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CTTGGAAGAT TCTGGAAGAA ACCTG 25

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CGCCAGGGTT TTCCCAGTCA CGAC 24

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GCCCCTTCTC CGGGAAAACG AATG 24

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GGAAACAGCT ATGACCATGA TTAC 24

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CAGGACTGTG CCCTTCAACC AGAC 24

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CGGTCTCAGG TTTCTTCCAG AATC 24

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CTGCAGGTAC CGGTCCGGAA TTCC 24

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CCATCTGTGA ATCTAACCTC AGTCTC 26

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

AACTCACATA ACATTTCCAG ATTG 24

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Asp Tyr Lys Asp Asp Asp Asp Lys
5

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3150 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 406..912

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 2693..2986

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

```
TATTCTATTC CAGCACAAGG CAGATGCACA GCAAATGTGA GCTGACTCTA GTCCTTCTTC      60

TGAAAACAGC CATGGGAAAT TTAGGCAAAG AATGTGTTGT CTTTGCTAAT ACTGCTCTTT     120

AAGCCCCAGA CATAGCTAAA CTCTTAGCTA ATTACCCCCT GGGTCCCAGG CTTTCACTGG     180

GGCCTTTTAA AATACACAAA ACCAAAGTGA CGGCAGGAGG CCATTAGCAC TACATAATTC     240

AAGCAAACAA TAAATGTGTT TATTCTGCCT GGCTACTGAC CACCTGCCTT CCCATCCCGC     300

CAGGCAGGTA TCTATATATA CGATTTCCTT TTTCCCAGTC CTGCAGAGAA TGAGCCTCTC     360

CTTTGGGCCT CATCATTTAC AAAAGAAGCT TGGGCCCCTG ACAGC ATG CAT CTC         414
                                                  Met His Leu
                                                    1

CTC TTA TTT CAG CTG CTG GTA CTC CTG CCT CTA GGA AAG ACC ACA CGG       462
Leu Leu Phe Gln Leu Leu Val Leu Leu Pro Leu Gly Lys Thr Thr Arg
      5                  10                  15

CAC CAG GAT GGC CGC CAG AAT CAG AGT TCT CTT TCC CCC GTA CTC CTG       510
His Gln Asp Gly Arg Gln Asn Gln Ser Ser Leu Ser Pro Val Leu Leu
 20                  25                  30                  35

CCA AGG AAT CAA AGA GAG CTT CCC ACA GGC AAC CAT GAG GAA GCT GAG       558
Pro Arg Asn Gln Arg Glu Leu Pro Thr Gly Asn His Glu Glu Ala Glu
                 40                  45                  50

GAG AAG CCA GAT CTG TTT GTC GCA GTG CCA CAC CTT GTA GCC ACC AGC       606
Glu Lys Pro Asp Leu Phe Val Ala Val Pro His Leu Val Ala Thr Ser
             55                  60                  65

CCT GCA GGG GAA GGC CAG AGG CAG AGA GAG AAG ATG CTG TCC AGA TTT       654
Pro Ala Gly Glu Gly Gln Arg Gln Arg Glu Lys Met Leu Ser Arg Phe
         70                  75                  80

GGC AGG TTC TGG AAG AAG CCT GAG AGA GAA ATG CAT CCA TCC AGG GAC       702
Gly Arg Phe Trp Lys Lys Pro Glu Arg Glu Met His Pro Ser Arg Asp
```

```
     85                  90                  95

TCA GAT AGT GAG CCC TTC CCA CCT GGG ACC CAG TCC CTC ATC CAG CCG      750
Ser Asp Ser Glu Pro Phe Pro Pro Gly Thr Gln Ser Leu Ile Gln Pro
100                 105                 110                 115

ATA GAT GGA ATG AAA ATG GAG AAA TCT CCT CTT CGG GAA GAA GCC AAG      798
Ile Asp Gly Met Lys Met Glu Lys Ser Pro Leu Arg Glu Glu Ala Lys
                120                 125                 130

AAA TTC TGG CAC CAC TTC ATG TTC AGA AAA ACT CCG GCT TCT CAG GGG      846
Lys Phe Trp His His Phe Met Phe Arg Lys Thr Pro Ala Ser Gln Gly
            135                 140                 145

GTC ATC TTG CCC ATC AAA AGC CAT GAA GTA CAT TGG GAG ACC TGC AGG      894
Val Ile Leu Pro Ile Lys Ser His Glu Val His Trp Glu Thr Cys Arg
        150                 155                 160

ACA GTG CCC TTC AGC CAG GTATGTGTTC TGGGGGGAGA GCAGGTAAGA             942
Thr Val Pro Phe Ser Gln
    165

GTTTGCAGGT GGTAGTGGAC AGCTGGGATG GATGGAGAGT AGGGGAAAAG GCTGTCAGGA    1002

GCCTGACTCT AGCTTAACTA CAGATTTGGT CCTTGGGCAT TCATCATAGG ATTTGGCAAA    1062

GATTAAGTTT CCTTCTGGCC TTTACCATTT TTTCTTGGCA TTGTGGAAAT GCTGCAAGAA    1122

TGATATGATG ATACTGTCAA TATCAGTAAT CATTCATTCA CACTGAAGAC ACAGAGCTCT    1182

GTTTTATTTA TTTATTTTTG CATTGGAGGT GATCTACTCA GAGATATAAG TCAGACTGTA    1242

CCCTCAGTTA-GGAAACTGAG AATTTAGAGT AATCACCAGA ACTCCTCTGT AGCTATCTTT    1302

CTGCACTCTA TTAATATGTG GATGAGCAGG TCAACTCCAT TTGTTGATAA AGTGGGGTGC    1362

ATTGGACTCC TTCCCAAATA CTCTCATATC CATTTACGAT GGTCTTAATC CCCATAGTCC    1422

ATACTTAATT ACTTTATAGG TTTATGAGGG ACTTCTTTAA TAGCTTGCTA AAGCTATCCC    1482

ACAACCTCAA AGTACGTTGA GGTTCTCAGG CAAAAGTTGT CATATCATTT CTAGTATTAT    1542

GATAGCAAAA AAGTGATTTT CTTTCACTTA TTTTCTCATA TGAGCTTTTT AAAAAATCAA    1602

TCTTGATGTG AGATCATATC TCCTCCCCTT AGAAGTACCT TTCTCCTGAT TCATGTTGTG    1662
```

```
TTGGCTGATT TGTAGTTATT ATGATCAATT CCATGCTATT AAGACAAAGG GACATCCTAC     1722

TGTCTACTTC CTCTGGCAAT ATCTACATTC CAAATGTTAA ATTAAAATTG AGAACTTGCA     1782

TTAGGTCCTT AAACATGAAG ATATTGAACC AAAAACATGC AGGGTAGAGT AAAATTTTAT     1842

AGTCGAGTAA TGCTACCCAA TTAAGCAAGC AATAGAATAG GGCAATTGAC TGTTCAAGGC     1902

AGTTAAGTAT TCTGCCTGAA AAGGCAAGGA TATGTAGCAA TGGCAAGTCA ATTATCAAAT     1962

AATAATGACT ACTCTGTTGG CCATGTGCAA TTAGAAAATT ACCCCTAAGA ATCAGGCAAT     2022

CAAATTTCTT TTGAAATTCT TCTTTTGAAT TCTATTGCTA ATTAAATTAA AACTAAGATG     2082

TTTGACTCTT ACATATTTTG AAAGGCATAT AAAGCTAGGT GCTTGGAGTT ATGAGAGGTA     2142

AAGGTGATGT AATATACAAT GATTTGCAGG CATATGCATT GTAACTCTGC TTGCATACAA     2202

CTTCATAGAC TTGAATGTAC TACAGGTCTT GCAGAATAGG ATAGAATTAA ACCTAGAATG     2262

TTCTGATCTA TTCTACGATC AATGTAACAA ATATGTATTG GGAGCCTACT ATGCACAAAG     2322

CCCTGTGAGG AATAAAAAAG TAAGGCACAT TACTTATGTA AGATAATTAC CATTAGAATT     2382

TTTCAATCGC TCACATCCAA TTAGACAAAA TTGCTTAAGG TTTTGCACGA ATAATGTAGA     2442

GTAAAATATT TTTTATGTTA ACTTAGGGAT TCCCTAAAGG CTGTTTAATA ATTTACTCAA     2502

TAAAGAAAAT TTAATTGAGG TGGTTCTGTG CCCTTATAGA TACCATCACT TGCATATTGC     2562

AAATTGTATC CAAAATTGGA AGCTTTGAA ATTTTTAAAT TATCCTCAGA TTTACAGTCC     2622

ATAGCTTCTG CATTATGTGT GTTAAAGAAA TAATTCAAAA TAACGTAATG GAAATGTGTT     2682

TGCTTTTTAG ACT ATA ACC CAC GAA GGC TGT GAA AAA GTA GTT GTT CAG         2731
            Thr Ile Thr His Glu Gly Cys Glu Lys Val Val Val Gln
              1               5                  10

AAC AAC CTT TGC TTT GGG AAA TGC GGG TCT GTT CAT TTT CCT GGA GCC       2779
Asn Asn Leu Cys Phe Gly Lys Cys Gly Ser Val His Phe Pro Gly Ala
     15                  20                  25

GCG CAG CAC TCC CAT ACC TCC TGC TCT CAC TGT TTG CCT GCC AAG TTC       2827
Ala Gln His Ser His Thr Ser Cys Ser His Cys Leu Pro Ala Lys Phe
 30                  35                  40                  45
```

```
ACC ACG ATG CAC TTG CCA CTG AAC TGC ACT GAA CTT TCC TCC GTG ATC      2875
Thr Thr Met His Leu Pro Leu Asn Cys Thr Glu Leu Ser Ser Val Ile
                 50                  55                  60

AAG GTG GTG ATG CTG GTG GAG GAG TGC CAG TGC AAG GTG AAG ACG GAG      2923
Lys Val Val Met Leu Val Glu Glu Cys Gln Cys Lys Val Lys Thr Glu
             65                  70                  75

CAT GAA GAT GGA CAC ATC CTA CAT GCT GGC TCC CAG GAT TCC TTT ATC      2971
His Glu Asp Gly His Ile Leu His Ala Gly Ser Gln Asp Ser Phe Ile
         80                  85                  90

CCA GGA GTT TCA GCT TGAAGAGCTA TCCCACTATT ACCTTTGAAA AGCAAAACCA      3026
Pro Gly Val Ser Ala
     95

CAACAGCAAA GATGCTGATT ATTCAGTCTG AAAATGTTAA GTGGGTACAT AACATTTTCA    3086

GGGAAAGGTG ACTTGAAACG TAGTTTTAAA TTAGAACGAT AGAGGAAATG ATATTAGTCT    3146

AGTT                                                                3150
```

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 169 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

```
Met His Leu Leu Leu Phe Gln Leu Leu Val Leu Leu Pro Leu Gly Lys
 1               5                  10                  15

Thr Thr Arg His Gln Asp Gly Arg Gln Asn Gln Ser Ser Leu Ser Pro
            20                  25                  30

Val Leu Leu Pro Arg Asn Gln Arg Glu Leu Pro Thr Gly Asn His Glu
        35                  40                  45

Glu Ala Glu Glu Lys Pro Asp Leu Phe Val Ala Val Pro His Leu Val
    50                  55                  60
```

```
Ala Thr Ser Pro Ala Gly Glu Gly Gln Arg Gln Arg Glu Lys Met Leu
 65                  70                  75                  80

Ser Arg Phe Gly Arg Phe Trp Lys Lys Pro Glu Arg Glu Met His Pro
                 85                  90                  95

Ser Arg Asp Ser Asp Ser Glu Pro Phe Pro Pro Gly Thr Gln Ser Leu
            100                 105                 110

Ile Gln Pro Ile Asp Gly Met Lys Met Glu Lys Ser Pro Leu Arg Glu
        115                 120                 125

Glu Ala Lys Lys Phe Trp His His Phe Met Phe Arg Lys Thr Pro Ala
    130                 135                 140

Ser Gln Gly Val Ile Leu Pro Ile Lys Ser His Glu Val His Trp Glu
145                 150                 155                 160

Thr Cys Arg Thr Val Pro Phe Ser Gln
                165
```

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 98 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

```
Thr Ile Thr His Glu Gly Cys Glu Lys Val Val Val Gln Asn Asn Leu
 1               5                  10                  15

Cys Phe Gly Lys Cys Gly Ser Val His Phe Pro Gly Ala Ala Gln His
            20                  25                  30

Ser His Thr Ser Cys Ser His Cys Leu Pro Ala Lys Phe Thr Thr Met
        35                  40                  45

His Leu Pro Leu Asn Cys Thr Glu Leu Ser Ser Val Ile Lys Val Val
    50                  55                  60
```

Met Leu Val Glu Glu Cys Gln Cys Lys Val Lys Thr Glu His Glu Asp
65 70 75 80

Gly His Ile Leu His Ala Gly Ser Gln Asp Ser Phe Ile Pro Gly Val
85 90 95

Ser Ala

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 508 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 66..434

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CTGCAGAGAA TGAGCCTCTC CTTTGGGCCT CATCATTTAC AAAAGAAGCT TGGGCCCCTG 60

ACAGC ATG CAT CTC CTC TTA TTT CAG CTG CTG GTA CTC CTG CCT CTA 107
Met His Leu Leu Leu Phe Gln Leu Leu Val Leu Leu Pro Leu
1 5 10

GGA AAG ACC ACA CGG CAC CAG GAT GGC CGC CAG ACT ATA ACC CAC GAA 155
Gly Lys Thr Thr Arg His Gln Asp Gly Arg Gln Thr Ile Thr His Glu
15 20 25 30

GGC TGT GAA AAA GTA GTT GTT CAG AAC AAC CTT TGC TTT GGG AAA TGC 203
Gly Cys Glu Lys Val Val Val Gln Asn Asn Leu Cys Phe Gly Lys Cys
35 40 45

GGG TCT GTT CAT TTT CCT GGA GCC GCG CAG CAC TCC CAT ACC TCC TGC 251
Gly Ser Val His Phe Pro Gly Ala Ala Gln His Ser His Thr Ser Cys
50 55 60

TCT CAC TGT TTG CCT GCC AAG TTC ACC ACG ATG CAC TTG CCA CTG AAC 299

```
Ser His Cys Leu Pro Ala Lys Phe Thr Thr Met His Leu Pro Leu Asn
          65                  70                  75

TGC ACT GAA CTT TCC TCC GTG ATC AAG GTG GTG ATG CTG GTG GAG GAG      347
Cys Thr Glu Leu Ser Ser Val Ile Lys Val Val Met Leu Val Glu Glu
     80                  85                  90

TGC CAG TGC AAG GTG AAG ACG GAG CAT GAA GAT GGA CAC ATC CTA CAT      395
Cys Gln Cys Lys Val Lys Thr Glu His Glu Asp Gly His Ile Leu His
 95                 100                 105                 110

GCT GGC TCC CAG GAT TCC TTT ATC CCA GGA GTT TCA GCT TGAAGAGCTA      444
Ala Gly Ser Gln Asp Ser Phe Ile Pro Gly Val Ser Ala
                115                 120

TCCCACTATT ACCTTTGAAA AGCAAAACCA CAACAGCAAA GATGCTGATT ATTCAGTCTG    504

AAAA                                                                508
```

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 123 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

```
Met His Leu Leu Leu Phe Gln Leu Leu Val Leu Leu Pro Leu Gly Lys
  1               5                  10                  15

Thr Thr Arg His Gln Asp Gly Arg Gln Thr Ile Thr His Glu Gly Cys
            20                  25                  30

Glu Lys Val Val Val Gln Asn Asn Leu Cys Phe Gly Lys Cys Gly Ser
        35                  40                  45

Val His Phe Pro Gly Ala Ala Gln His Ser His Thr Ser Cys Ser His
    50                  55                  60

Cys Leu Pro Ala Lys Phe Thr Thr Met His Leu Pro Leu Asn Cys Thr
65                  70                  75                  80
```

Glu Leu Ser Ser Val Ile Lys Val Val Met Leu Val Glu Glu Cys Gln
85 90 95

Cys Lys Val Lys Thr Glu His Glu Asp Gly His Ile Leu His Ala Gly
100 105 110

Ser Gln Asp Ser Phe Ile Pro Gly Val Ser Ala
115 120

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Asp Leu His Val Ala Asn His Glu Glu Ala Glu Asp Lys Pro Asp Leu Phe Val Ala
5 10 15
Met Pro His Leu Met Gly
20 25

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Ala Pro Pro Pro Ile Asn Lys Leu Ala Leu Phe Pro Asp Lys Ser Ala Trp Cys Glu
5 10 15
Ala Lys Asn Ile Thr Gln Ile Val Gly His Ser
20 25 30

CLAIMS:

1. An isolated polypeptide of mammalian origin comprising a signal sequence and a domain conforming to the criteria for a cystine knot and optionally a long N-terminal domain between said signal sequence and cystine knot domain or a derivative of said polypeptide.

2. An isolated polypeptide according to claim 1 comprising the amino acid sequence:

C{AA}Q{AA}H{AA}C{AA}[X^1]$_n$QN{AA}C{AA}G{AA}C{AA}S{AA}P

wherein

{AA} is an amino acid sequence comprising from about 0 to about 50 amino acid residues;

X^1 is V or I; and

n is 0 or 1.

3. An isolated polypeptide according to claim 2 wherein said polypeptide exhibits the following characteristics:

(i) being glycosylated in its naturally occurring form;

(ii) being secretable in its naturally occurring form;

(iii) comprising a signal sequence and a domain conforming to the criteria for a cystine knot,

said cystine knot domain comprising the sequence:

CxTxPFxQxIxHExCxxxVxQNNLCFGKCxSxxxPx$_{n1}$CSHCxP [SEQ ID NO:1]

wherein x is any amino acid residue and n_1 is from about 6 to about 10 or comprises a sequence in the cystine knot domain having at least 50% identity to SEQ ID NO:1 excluding the cystine and x residues.

4. An isolated polypeptide according to claim 3 wherein the cystine knot domain

comprises the amino acid sequence:

CRTVPFNQTIAHEDCQKVVVQNNLCFGKC [SEQ ID NO:2]

or a sequence having at least 45% similarity to SEQ ID NO:2.

5. An isolated polypeptide according to any one of claims 1 to 4 comprising an amino acid sequence having at least 20% similarity to cerberus protein from *Xenopus laevis* as defined in Figure 1.

6. An isolated polypeptide according to any one of claims 1 to 4 comprising a sequence of amino acids substantially as set forth in SEQ ID NO:4 or having at least 50% similarity thereto and which polypeptide has the identifying characteristics of a CRP cytokine.

7. An isolated polypeptide according to any one of claims 1 to 4 comprising a sequence of amino acids substantially as set forth in SEQ ID NO:5 or having at least 50% similarity thereto and which polypeptide has the identifying characteristics of a CRP cytokine.

8. An isolated polypeptide according to any one of claims 1 to 4 comprising a sequence of amino acids substantially as set forth in SEQ ID NO:6 or having at least 50% similarity thereto and which polypeptide has the identifying characteristics of a CRP cytokine.

9. An isolated polypeptide according to any one of claims 1 to 4 comprising a sequence of amino acids substantially as set forth in SEQ ID NO:7 or having at least 50% similarity thereto and which polypeptide has the identifying characteristics of a CRP cytokine.

10. An isolated polypeptide according to any one of claims 1 to 4 comprising a sequence of amino acids substantially as set forth in SEQ ID NO:19 and/or 20 or having at least 50% similarity thereto and which polypeptide has the identifying characteristics of a CRP cytokine.

11. An isolated polypeptide according to any one of claims 1 to 4 comprising a sequence

of amino acids substantially as set forth in SEQ ID NO:22 or having at least 50% similarity thereto and which polypeptide has the identifying characteristics of a CRP cytokine.

12. An isolated nucleic acid molecule having a nucleotide sequence encoding a signal sequence and a domain conforming to the criteria for a cystine knot and optionally a long N-terminal domain between said signal sequence and cystine knot domain or a derivative of said polypeptide.

13. An isolated nucleotide acid molecule according to claim 12 encoding the amino acid sequence:

$$C\{AA\}Q\{AA\}H\{AA\}C\{AA\}[X^1]_nQN\{AA\}C\{AA\}G\{AA\}C\{AA\}S\{AA\}P$$

wherein

{AA} is an amino acid sequence comprising from about 0 to about 50 amino acid residues;

X^1 is V or I; and

n is 0 or 1.

14. An isolated nucleic acid molecule according to claim 13 encoding a polypeptide exhibiting the following characteristics:

(i) being glycosylated in its naturally occurring form;

(ii) being secretable in its naturally occurring form;

(iii) comprising a signal sequence and a domain conforming to the criteria for a cystine knot, said cystine knot domain comprising the sequence:

$$CxTxPFxQxIxHExCxxxVxQNNLCFGKCxSxxxPx_{n1}CSHCxP \text{ [SEQ ID NO:1]}$$

wherein x is any amino acid residue and n_1 is from about 6 to about 10 or comprises a sequence in the cystine knot domain having at least 50% identity to SEQ ID NO:1 excluding the cystine and x residues.

15. An isolated nucleic acid molecule according to claim 14 encoding a polypeptide comprising a cystine knot domain comprises the amino acid sequence:

CRTVPFNQTIAHEDCQKVVVQNNLCFGKC [SEQ ID NO:2]

or a sequence having at least 65% similarity to SEQ ID NO:2.

16. An isolated nucleic acid molecule according to any one of claims 12 to 15 encoding a polypeptide with an amino acid sequence having at least 20% homology to cerberus protein from *Xenopus laevis* as defined in Figure 1.

17. An isolated nucleic acid molecule according to any one of claims 12 to 15 encoding a polypeptide with a sequence of amino acids substantially as set forth in SEQ ID NO:4 or having at least 50% similarity thereto and which polypeptide has the identifying characteristics of a CRP cytokine.

18. An isolated nucleic acid molecule according to any one of claims 12 to 15 encoding a polypeptide with a sequence of amino acids substantially as set forth in SEQ ID NO:5 or having at least 50% similarity thereto and which polypeptide has the identifying characteristics of a CRP cytokine.

19. An isolated nucleic acid molecule according to any one of claims 12 to 15 encoding a polypeptide with a sequence of amino acids substantially as set forth in SEQ ID NO:6 or having at least 50% similarity thereto and which polypeptide has the identifying characteristics of a CRP cytokine.

20. An isolated nucleic acid molecule according to any one of claims 12 to 15 encoding a polypeptide with sequence of amino acids substantially as set forth in SEQ ID NO:7 or having at least 50% similarity thereto and which polypeptide has the identifying characteristics of a CRP cytokine.

21. An isolated nucleic acid molecule according to any one of claims 12 to 15 encoding a polypeptide with a sequence of amino acids substantially as set forth in SEQ ID NO:19 and/or 20 or having at least 50% similarity thereto and which polypeptide has the identifying characteristics of a CRP cytokine.

22. An isolated nucleic acid according to any one of claims 12 to 15 encoding a polypeptide with a sequence of amino acids substantially as set forth in SEQ ID NO:22 or having at least 50% similarity thereto and which polypeptide has the identifying characteristics of a CRP cytokine.

23. A method for modulating expression of a CRP cytokine in a mammal, said method comprising contacting a gene encoding said CRP cytokine with an effective amount of a modulator of CRP cytokine expression for a time and under conditions sufficient to up-regulate or down-regulate or otherwise modulate expression of the CRP cytokine.

24. A method for modulating activity of the CRP cytokine in a mammalian, said method comprising administering to said mammal a modulating effective amount of a molecule for a time and under conditions sufficient to increase or decrease CRP cytokine activity. The molecule may be a proteinaceous molecule or a chemical entity and may also be a derivative of a CRP cytokine or its ligand or a chemical analogue or truncation mutant of a CRP cytokine or its ligand.

25. A composition comprising one or more CRP cytokines or derivatives thereof or a modulator of CRP cytokine expression or CRP cytokine activity and one or more pharmaceutically acceptable carriers and/or diluents.

26. A method for detecting CRP in a biological sample from a subject said method comprising contacting said biological sample with an antibody specific for a CRP or group of CRPs or their derivatives or homologues for a time and under conditions sufficient for an antibody-CRP complex to form, and then detecting said complex.

27. Use of a CRP cytokine or its functional derivatives in the manufacture of a medicament for the treatment of defective or deficient CRP mediated activities.

FIG 1A

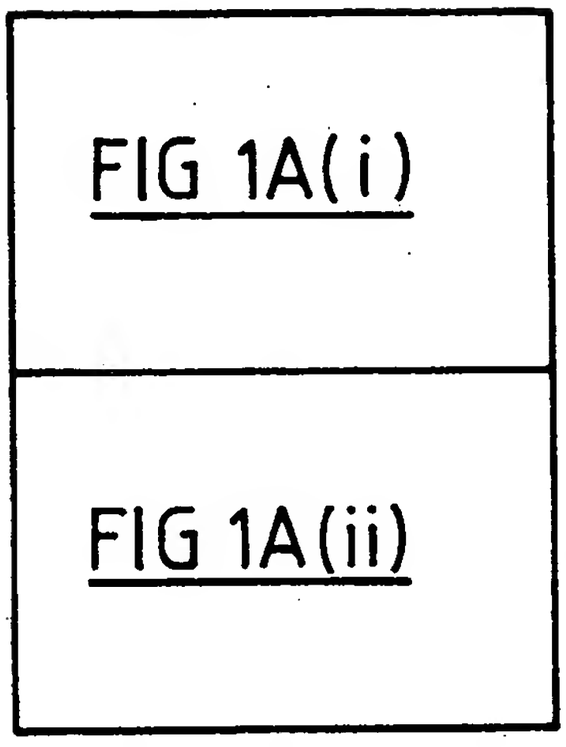


Figure 1A (i)

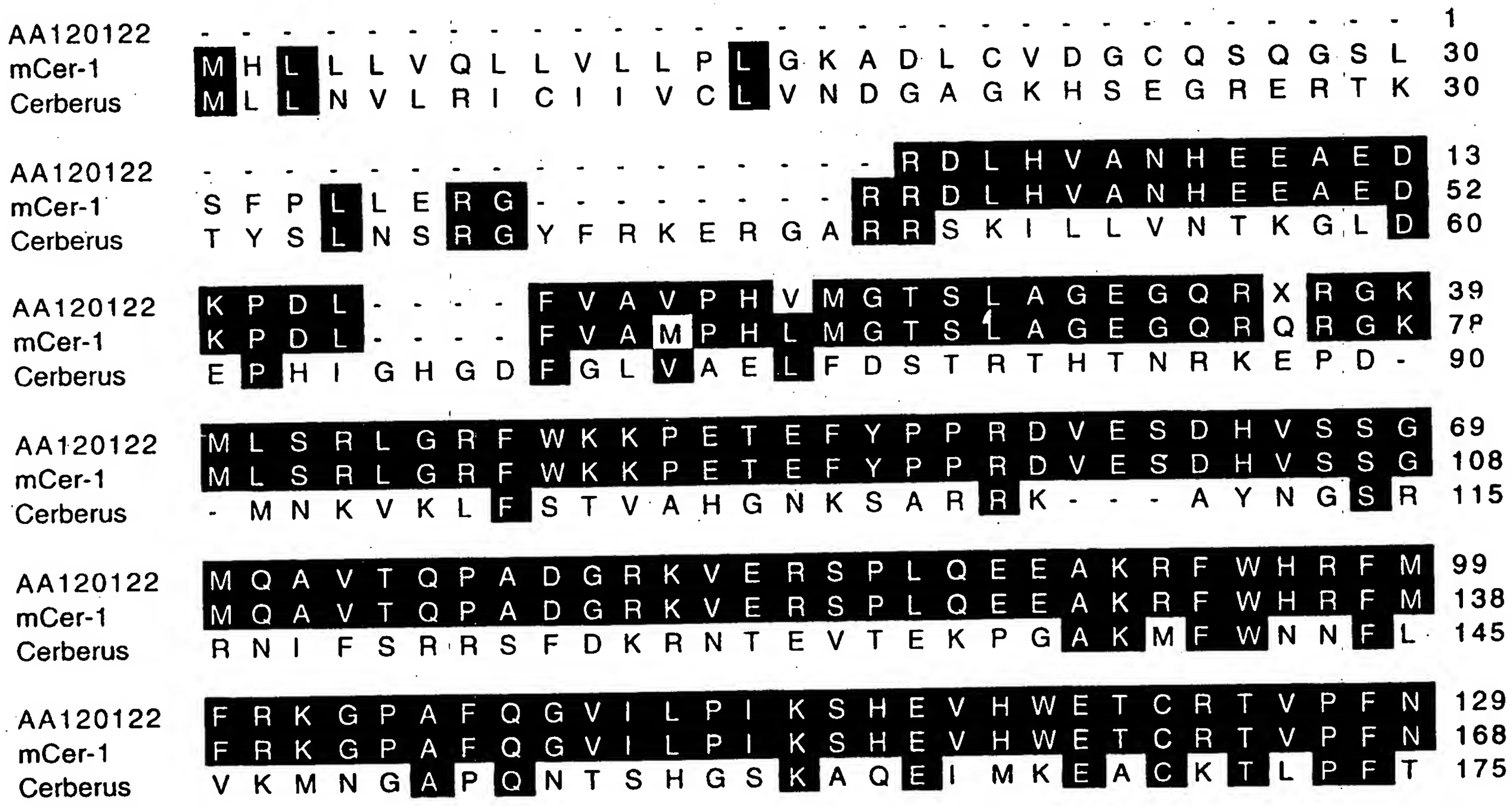

AA120122 - 1
mCer-1 M H L L L V Q L L V L L P L G K A D L C V D G C Q S Q G S L 30
Cerberus M L L N V L R I C I I V C L V N D G A G K H S E G R E R T K 30
AA120122 - - - - - - - - - - - - - - - - - R D L H V A N H E E A E D 13
mCer-1 S F P L L E R G - - - - - - - - - R R D L H V A N H E E A E D 52
Cerberus T Y S L N S R G Y F R K E R G A R R S K I L L V N T K G L D 60
AA120122 K P D L - - - - F V A V P H V M G T S L A G E G Q R X R G K 39
mCer-1 K P D L - - - - F V A M P H L M G T S L A G E G Q R Q R G K 79
Cerberus E P H I G H G D F G L V A E L F D S T R T H T N R K E P D - 90
AA120122 M L S R L G R F W K K P E T E F Y P P R D V E S D H V S S G 69
mCer-1 M L S R L G R F W K K P E T E F Y P P R D V E S D H V S S G 108
Cerberus - M N K V K L F S T V A H G N K S A R R K - - - A Y N G S R 115
AA120122 M Q A V T Q P A D G R K V E R S P L Q E E A K R F W H R F M 99
mCer-1 M Q A V T Q P A D G R K V E R S P L Q E E A K R F W H R F M 138
Cerberus R N I F S R R S F D K R N T E V T E K P G A K M F W N N F L 145
AA120122 F R K G P A F Q G V I L P I K S H E V H W E T C R T V P F N 129
mCer-1 F R K G P A F Q G V I L P I K S H E V H W E T C R T V P F N 168
Cerberus V K M N G A P Q N T S H G S K A Q E I M K E A C K T L P F T 175

Figure 1A (ii)

```
AA120122   Q T I A H E D C Q K V V V Q N N L C F G K C S S I R F P G E 159
mCer-1     Q T I A H E D C Q K V V V Q N N L C F G K C S S I R F P G E 198
Cerberus   Q N I V H E N C D R M V I Q N N L C F G K C I S L H V P N Q 205

AA120122   G A - - - - - - - - - - - - - - - - - - - - - - - - - - - 162
mCer-1     G A D A H S F C S H C S P T K F T T V H L M L N C T S P T P 228
Cerberus   - Q D R R N T C S H C L P S K F T L N H L T L N C T G S K N 234

AA120122   - - - - - - - - - - - - - - - - - - - - - - - - - - - - - 162
mCer-1     V V K M V M Q V E E C Q C M V K T E R G E E R L L L A G S Q 258
Cerberus   V V K V V M M V E E C T C E A H K S N F H Q T - - - - - - - 258

AA120122   - - - - - - - - - - - - - - - - D                         162
mCer-1     G S F I P G L P A S K T N P .                             273
Cerberus   A Q F N M D - T S T T L H H .                             271
```

PCT/AU98/00078

```
mCer-1     MHLLLVQLLVLLPLGKADLCVDGCQSQGSLSFPLLENG--  39
Cerberus   MLLNVLRICIIVCLVNDGAGKHSEGRERTKTYSLNSFFYF  40
DAN        MLWVLVGAVLPVMLLAAPPP----------------------  21

mCer-1     -------RRDLHVANHEEAEDKPDL----FVAMPHLMGTSL  68
Cerberus   RKERGARRSKILLVNTKGLDEFHIGHGDFGLVAELFDSTR  80
DAN        ----------------------------------------  21

mCer-1     AGEGQRQRGKMLSRLGRFWKKPETEFYPPRDVESDHVSSG  108
Cerberus   THTNRKEPD--MNKVKLFSTVAHGNKSARRKA---YNGSR  115
DAN        ----------INKLALFPDKSAW-----------------  34

mCer-1     MQAVTQPADGRKVERSPLQEEAKRFWHRFMFRKGPAFQGV  148
Cerberus   RNIFSRRSFDKRNTEVTEKPGAKMFWNNFLVKMNGAPQNT  155
DAN        ----------------------------------------  34

mCer-1     ILPIKSHEVHWETCRTVPFNQTIAHEDCQKVVVQNNLCFG  188
Cerberus   SHGSKAQEIMKEACKTLPFTQNIVHENCDRMVIQNNLCFG  195
DAN        ----------------CEAKNITQIVGHSGCEAKSIQNRACLG  60

mCer-1     KCSSIRFPG---EGADAHSFCSHCSFTKFTTVHLMLNCTS  225
Cerberus   KCISLHVPN---Q-QDRRNTCSHCLPSKFTLNHLTLECTG  231
DAN        QCFSYSVPNTFPQSTESLVHCDSCMPAQSMWEIVTLECPD  100

mCer-1     PTPVVKMVMQVE---ECQCMV-KTERGEERLLL--------  255
Cerberus   SKNVVKVVMMVE---ECTCEA-HKSNFHQTA----------  259
DAN        HEEVPRVDKLVEKIVHCSCQACGKEPSHEGLNVYVQGEDS  140

mCer-1     AGSQG-----------------------------SFIPGLPASKTNP.  273
Cerberus   ----------------------------------QFNMD-TSTTLHH.  271
DAN        PGSQPGPHSHAHPHPGGQTPEPEEPPGAPQVEEEGAED.  179
```

Figure 1B

Figure 1C

NDP	M R N H V L A A S I S M L S L L A I M G D T D S K T D S S F	30
DAN	M L W V L V G A V L P V M L L A A P P P I N K - - - L A L F	27
NDP	L M D S Q R C M R H H Y V D S I S H P L Y K C S S K M V L L	60
DAN	P D K S A W C E A K N I T Q I V G H S - - G C E A K S I Q N	55
NDP	A R C E G H C S Q A S R S E P L V S F S T V L K Q P F R S S	90
DAN	R A C L G Q C F S Y S V P N T F P Q S T E S L V H - - - - -	81
NDP	C H C C R P Q T S K L K A L R L R C S G - - - G M R L T A T	117
DAN	C D S C M P A Q S M W E I V T L E C P D H E E V P R V D K L	110
NDP	Y R Y I L S C H C E E C S S .	132
DAN	V E K I V H C S C Q A C G K E P S H E G L N V Y V Q G E D S	140
NDP		132
DAN	P G S Q P G P H S H A H P H P G G Q T P E P E E P P G A P Q	170
NDP		132
DAN	V E E E G A E D .	179

Figure 1D

```
mCer-1  MHLLLVQLLVLLPLGKADLCVDGCQSQGSLSFPLLERGRR  40
DAN     MLWVLVGAVLPVMLLAAPPPINKLA-----------------  26

mCer-1  DLHVANHEEAEDKPDLFVAMPHLMGTSLAGEGQRQRGKML  80
DAN     ----------------------------------------  26

mCer-1  SRLGRFWKKPETEFYPPRDVESDHVSSGMQAVTQPADGRK  120
DAN     ----------------------------------------  26

mCer-1  VERSPLQEEAKRFWHRFMFRKGPAFQGVILPIKSHEVHWE  160
DAN     ----------------------------LFPDKSA      32

mCer-1  TCRTVPFNQTIAHEDCQKVVVQNNLCFGKCSSIRFPG---  198
DAN     WCEAKNITQIVGHSGCEAKSIQNRACLGQCFSYSVPNTFP  72

mCer-1  EGADAHSFCSHCSPTKFTTVHLMLNCTSPTPVV---KMVM  234
DAN     QSTESLVHCDSCMPAQSMWEIVTLECPDHEEVPRVDKLVE  112

mCer-1  QVEECQCMV-KTERGEERL--LLAGSQGSFIPGLPASKTN  271
DAN     KIVHCSCQACGKEPSHEGLNVYVQGEDSPGSQPGPHSHAH  152

mCer-1  P.                                        273
DAN     PHPGGQTPEPEEPPGAPQVEEEGAED.               179
```

Figure 1E

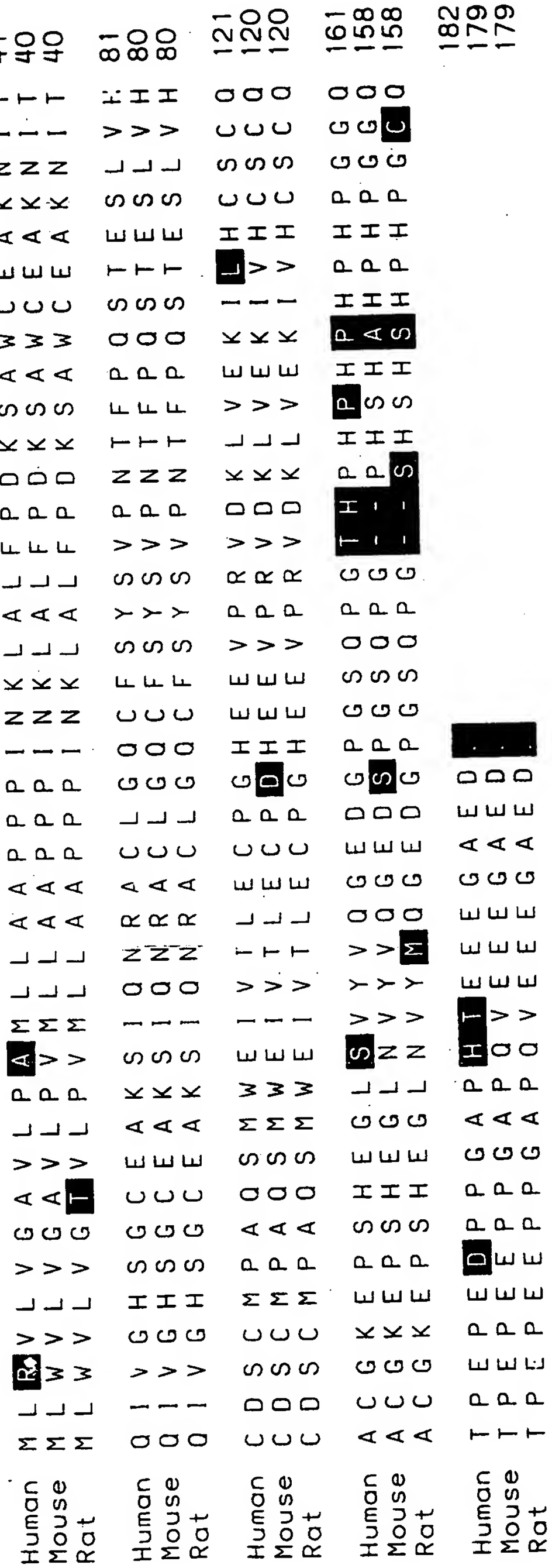

Human MLAAPPPPIKNKSSDKSS 41
Mouse MLAAPPPPIKNKSSDKSS 40
Rat MLAAPPPPIKNKSSDKSS 40
Human QIVGHSSGCEEAKKPDKK 81
Mouse QIVGHSSGCEEAKKPDKK 80
Rat QIVGHSSGCEEAKKPDKK 80
Human CDSCMPPAQSMWKKPDKK 121
Mouse CDSCMPPAQSMWKKPDKK 120
Rat CDSCMPPAQSMWKKPDKK 120
Human ACGKKPEEGSHKLLPDKK 161
Mouse ACGKKPEEGSHKLLPDKK 158
Rat ACGKKPEEGSHKLLPDKK 158
Human TPPEEPPPGSHKLLPDKK 182
Mouse TPPEEPPPGSHKLLPDKK 179
Rat TPPEEPPPGSHKLLPDKK 179

Figure 1F

```
Cerberus  M L L N V L R I C I I V C L V N D G A G K H S E G R E R T K T Y S L N S R G Y F  40
mCer-1    M H L L L V Q L L V L L P L G K A D L C V D G C Q S Q G S L S F P L L E R G - -  39

Cerberus  R K E R G A R R S K I L L V N T K G L D E P H I G H G D F G L V A E L F D S T R  80
mCer-1    - - - - - - R R D L H V A N H E E A E D K P D L - - - - F V A M P H L M G T S L  68

Cerberus  T H T N R K E P D - - M N K V K L F S T V A H G N K S A R R K A - - - Y N G S R 115
mCer-1    A G E G Q R Q R G K M L S R L G R F W K K P E T E F Y P P R D V E S D H V S S G 108

Cerberus  R N I F S R R S F D K R N T E V T E K P G A K M F W N N F L V K M N G A P Q N T 155
mCer-1    M Q A V T Q P A D G R K V E R S P L Q E E A K R F W H R F M F R K G P A F Q G V 148

Cerberus  S H G S K A Q E I M K E A C K T L P F T Q N I V H E N C D R M V I Q N N L C F G 195
mCer-1    I L P I K S H E V H W E T C R T V P F N Q T I A H E D C Q K V V V Q N N L C F G 188

Cerberus  K C I S L H V P N Q - Q D R R N T C S H C L P S K F T L N H L T L N C T G S K N 234
mCer-1    K C S S I R F P G E G A D A H S F C S H C S P T K F T T V H L M L N C T S P T P 228

Cerberus  V V K V V M M V E E C T C E A H K S N F H Q T - - - - - - - - A Q F N M D - T S T 266
mCer-1    V V K M V M Q V E E C Q C M V K T E R G E E R L L L A G S Q G S F I P G L P A S 268

Cerberus  T L H H .                                                                    271
mCer-1    K T N P .                                                                    273
```

Figure 2

```
TTTTAGGCCCGTCCATCTGTGAATCTAACCTCAGTCTCTGGGAATCAGGAAGCATGCATCTCCTCTTAGTTCAGCTGCTTGTTCTCTTGCCTCTGGGGAA 100
                                                        M  H  L  L  L  V  Q  L  L  V  L  L  P  L  G  K

GGCAGACCTATGTGTGGATGGCTGCCAGAGTCAGGGCTCTTTATCCTTTCCTCTCCTAGAAAGGGGTCGCAGAGATCTCCACGTGGCCAACCACGAGGAG 200
  A  D  L  C  V  D  G  C  Q  S  Q  G  S  L  S  F  P  L  L  E  R  G  R  R  D  L  H  V  A  N  H  E  E

GCAGAAGACAAGCCGGATCTGTTTGTGGCCATGCCACACCTCATGGGCACCAGCCTGGCTGGGGAAGGTCAGAGGCAGAGAGGGAAGATGCTGTCCAGGC 300
 A  E  D  K  P  D  L  F  V  A  M  P  H  L  M  G  T  S  L  A  G  E  G  Q  R  Q  R  G  K  M  L  S  R

TTGGAAGATTCTGGAAGAAACCTGAGACCGAATTTTACCCCCCAAGGGATGTGGAAAGCGATCATGTCTCATCGGGGATGCAGGCCGTGACTCAGCCAGC 400
L  G  R  F  W  K  K  P  E  T  E  F  Y  P  P  R  D  V  E  S  D  H  V  S  S  G  M  Q  A  V  T  Q  P  A

AGATGGGAGGAAAGTGGAGAGATCACCTCTACAGGAGGAAGCCAAGAGGTTCTGGCATCGGTTCATGTTCAGAAAGGGCCCGGCGTTCCAGGGAGTCATC 500
  D  G  R  K  V  E  R  S  P  L  Q  E  E  A  K  R  F  W  H  R  F  M  F  R  K  G  P  A  F  Q  G  V  I

CTGCCCATCAAAAGCCACGAAGTACACTGGGAGACCTGCAGGACTGTGCCCTTCAACCAGACCATTGCCCATGAAGACTGTCAAAAAGTCGTTGTCCAGA 600
 L  P  I  K  S  H  E  V  H  W  E  T  C  R  T  V  P  F  N  Q  T  I  A  H  E  D  C  Q  K  V  V  V  Q

ACAACCTTTGCTTTGGCAAATGCAGTTCCATTCGTTTTCCCGGAGAAGGGGCAGATGCCCACAGCTTCTGCTCCCACTGCTCGCCCACCAAATTCACCAC 700
N  N  L  C  F  G  K  C  S  S  I  R  F  P  G  E  G  A  D  A  H  S  F  C  S  H  C  S  P  T  K  F  T  T

CGTGCACTTGATGCTGAACTGCACCAGCCCAACCCCCGTGGTCAAGATGGTGATGCAAGTAGAAGAGTGTCAGTGCATGGTGAAGACGGAACGTGGAGAG 800
  V  H  L  M  L  N  C  T  S  P  T  P  V  V  K  M  V  M  Q  V  E  E  C  Q  C  M  V  K  T  E  R  G  E

GAGCGCCTCCTACTGGCTGGTTCCCAGGGTTCCTTCATCCCTGGACTTCCAGCTTCAAAAACAAACCCATGAATTACCTCAACAGAAAGCAAAACCTCAA 900
 E  R  L  L  L  A  G  S  Q  G  S  F  I  P  G  L  P  A  S  K  T  N  P  .

CAGAATAGGTGARGGTTATTCAATCTGGAAATGTTATGTGAGTTTATATAAAGATCAGTGGAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGG 1000
```

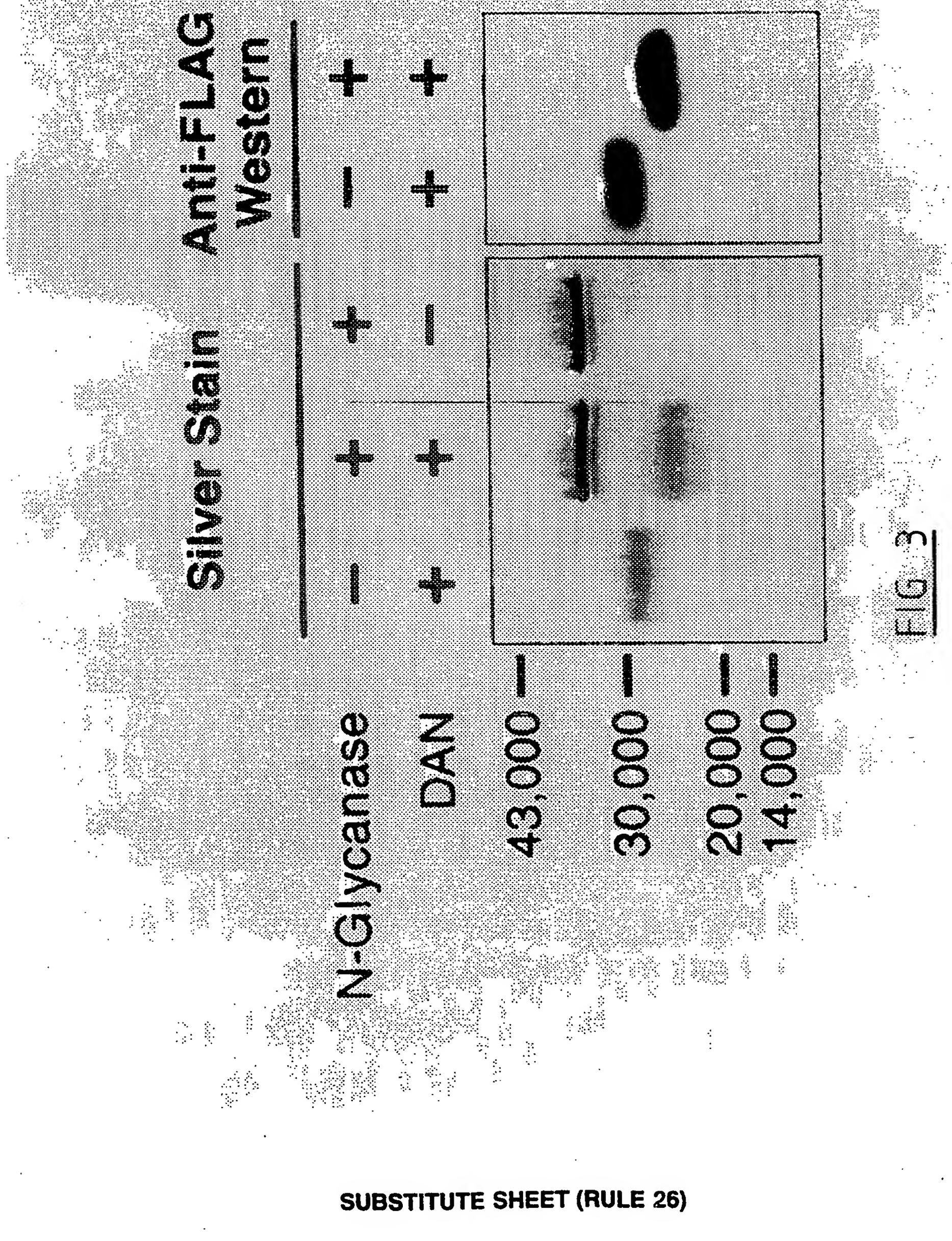
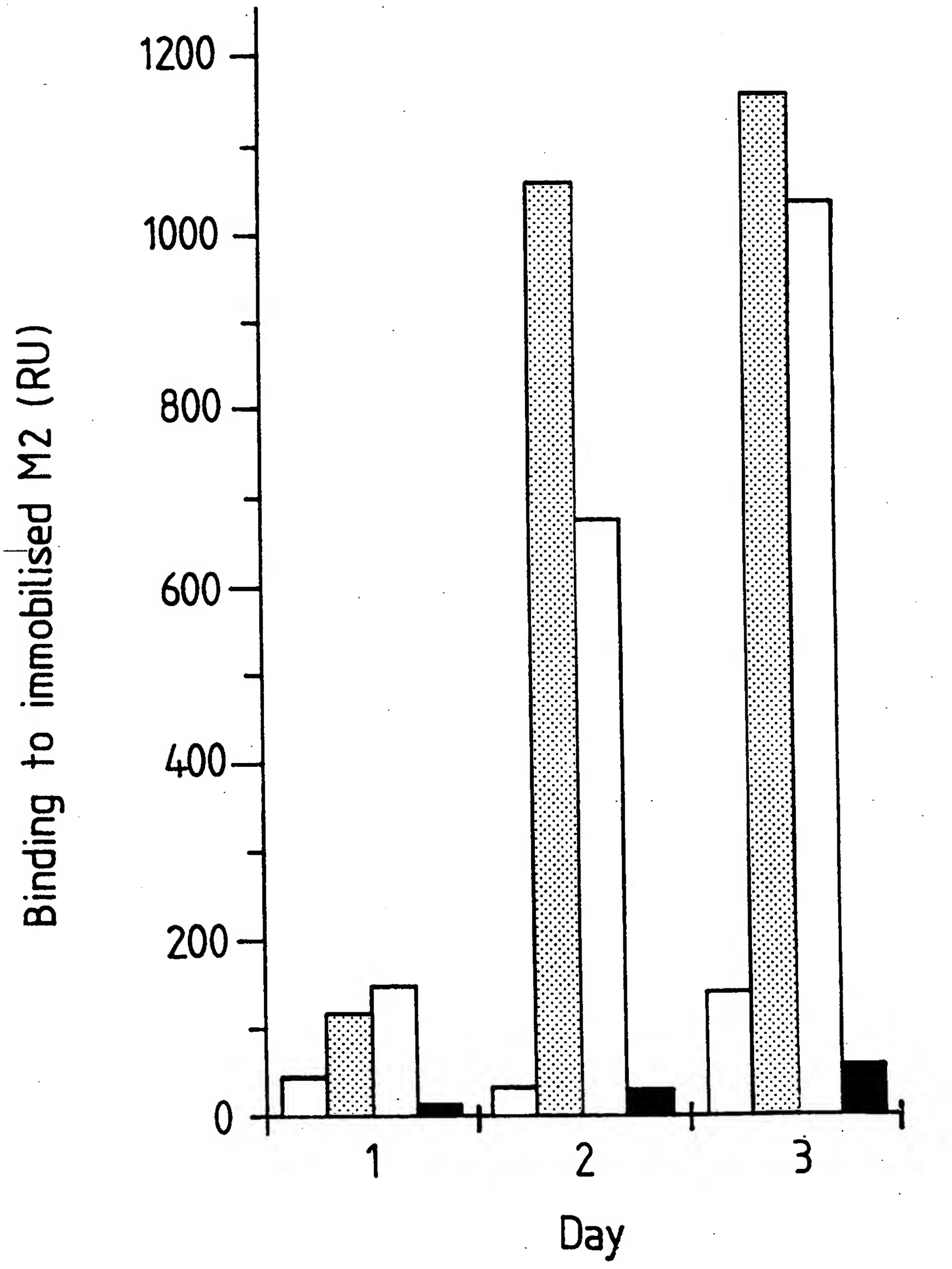

Silver Stain
Anti-FLAG Western
N-Glycanase
DAN
+ +
- +
+ -
+ +
- +
43,000
30,000
20,000
14,000

FIG 3

FIG 4


Binding to immobilised M2 (RU)
1200
1000
800
600
400
200
0
1
2
3
Day

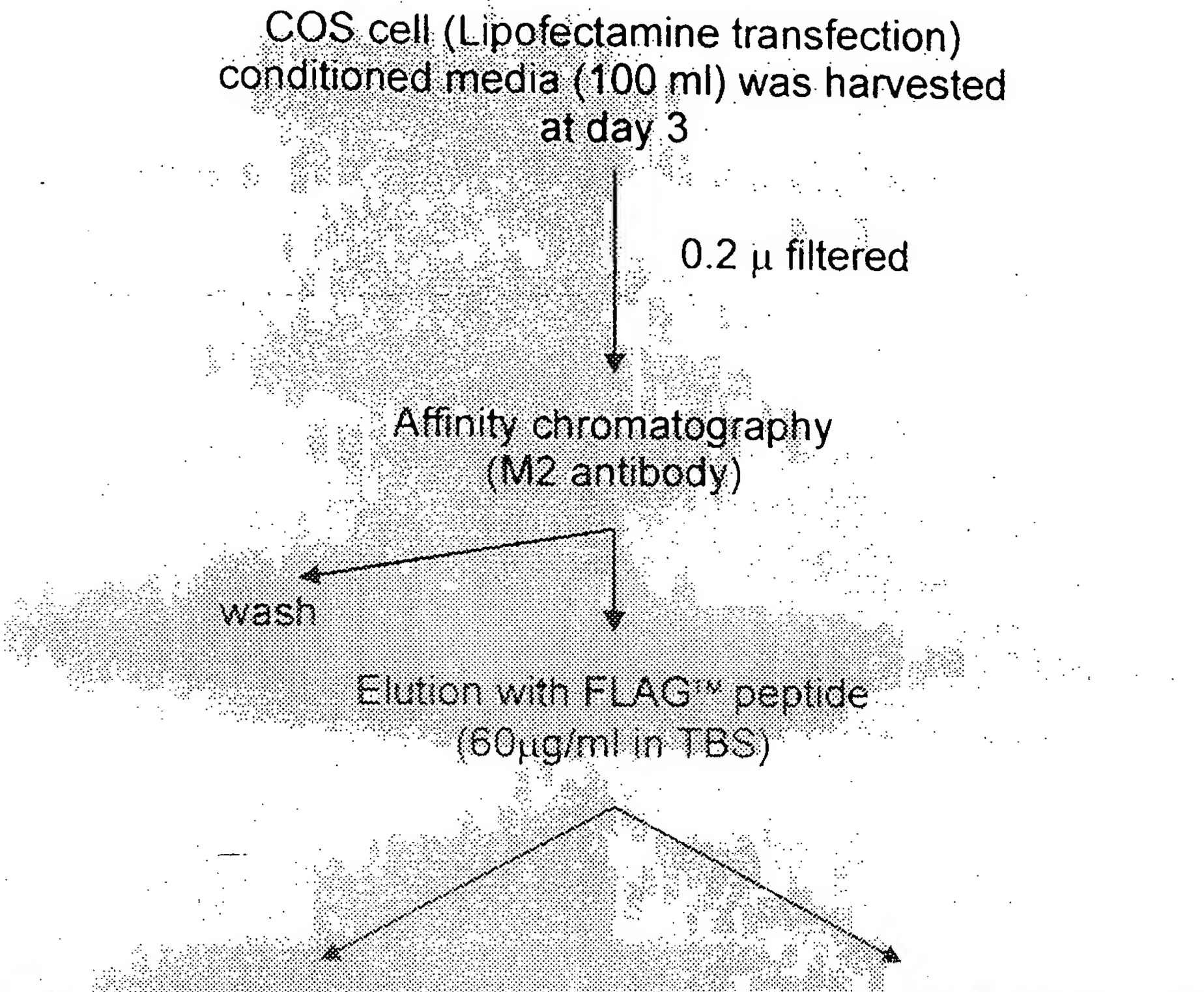

COS cell (Lipofectamine transfection)
conditioned media (100 ml) was harvested
at day 3
0.2 μ filtered
Affinity chromatography
(M2 antibody)
wash
Elution with FLAG™ peptide
(60μg/ml in TBS)
SDS-PAGE
(reducing)
93
67
43
30
20
14
1 2 3 4 5
Western blot
(reducing)
206
114
34
29
20
7
1 2 3 4 5

FIG 5

FIG 6

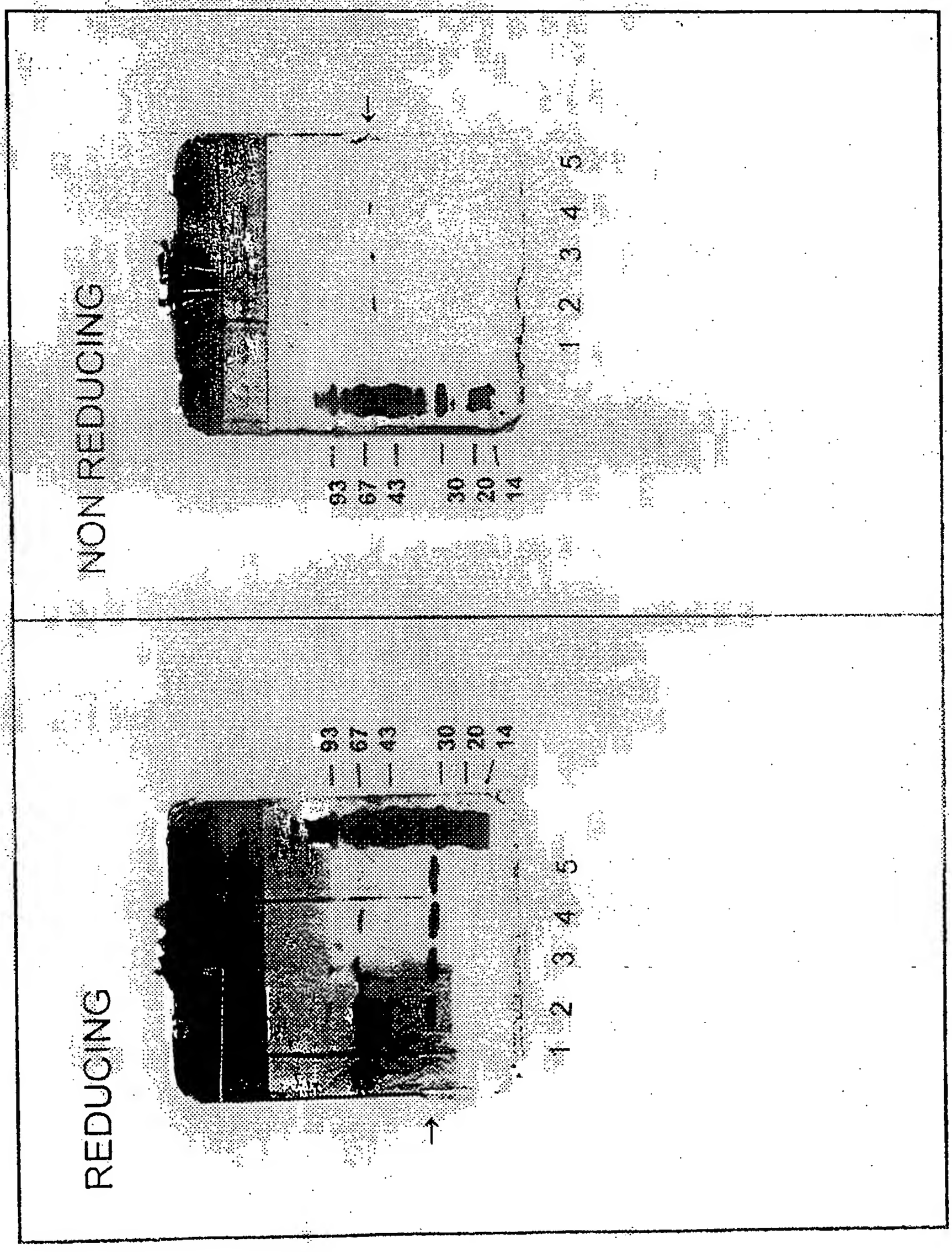
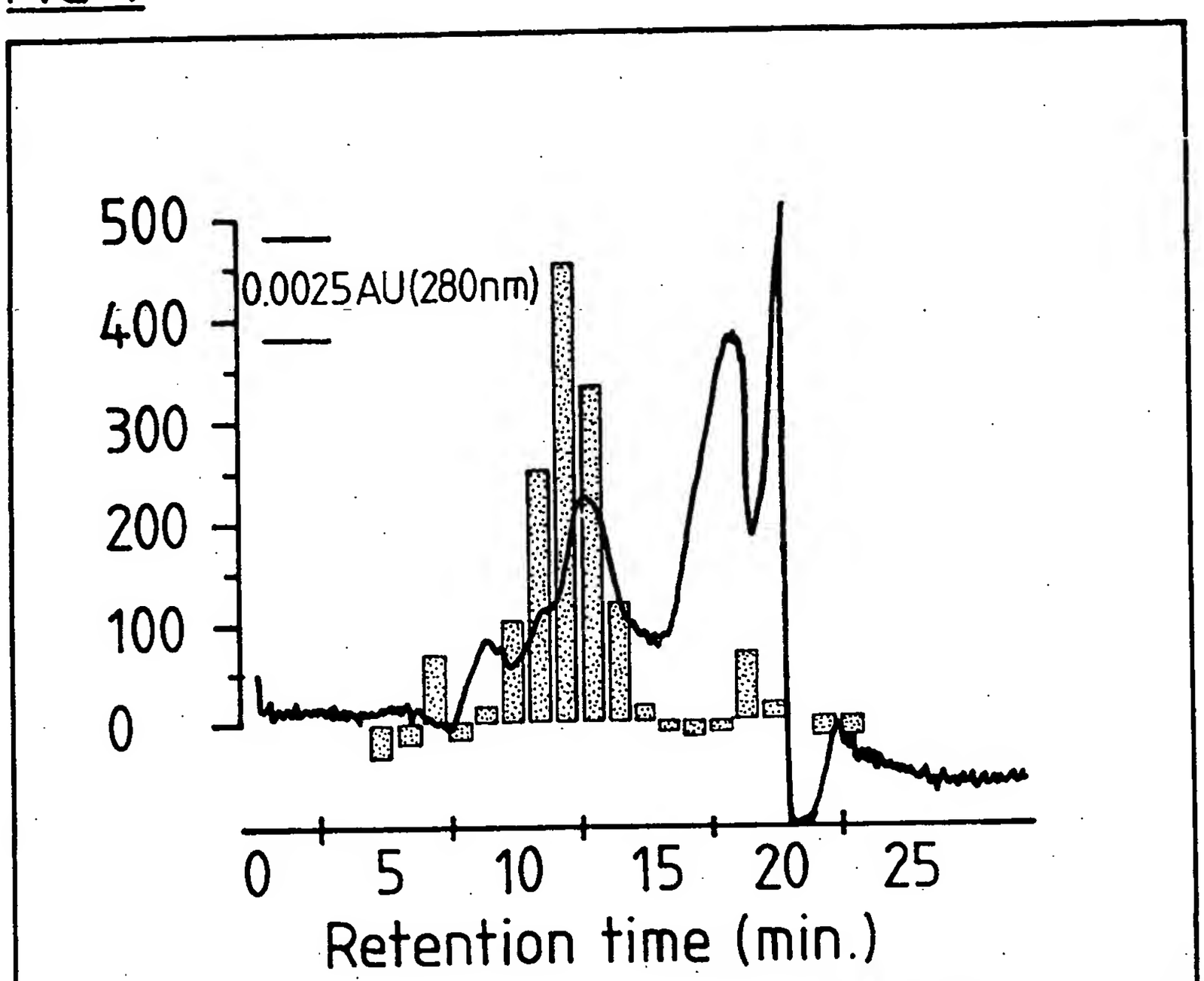
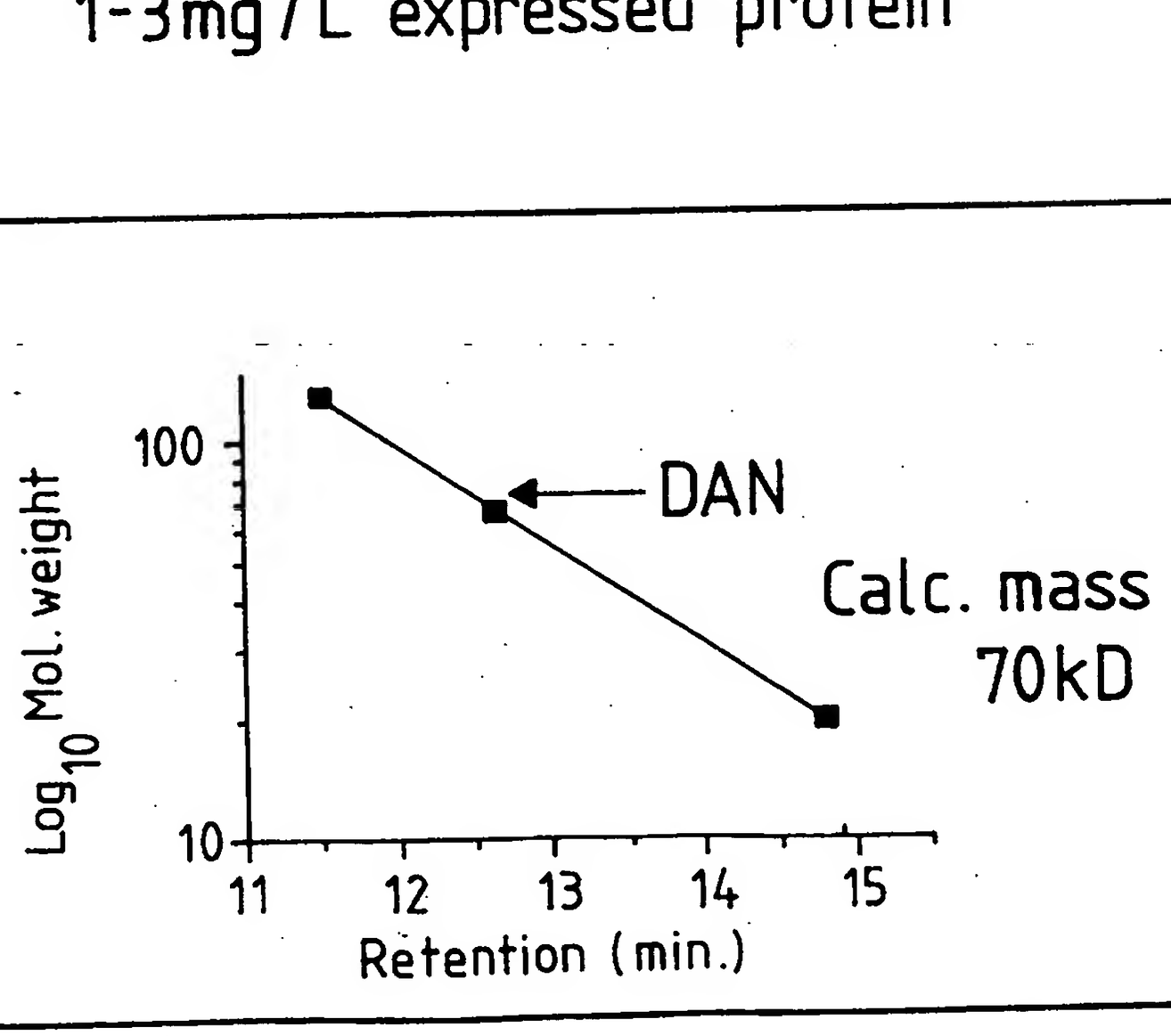

NON REDUCING
63
67
43
30
29
14
1 2 3 4 5
REDUCING
63
67
43
30
29
14
1 2 3 4 5

FIG 7



1-3mg/L expressed protein



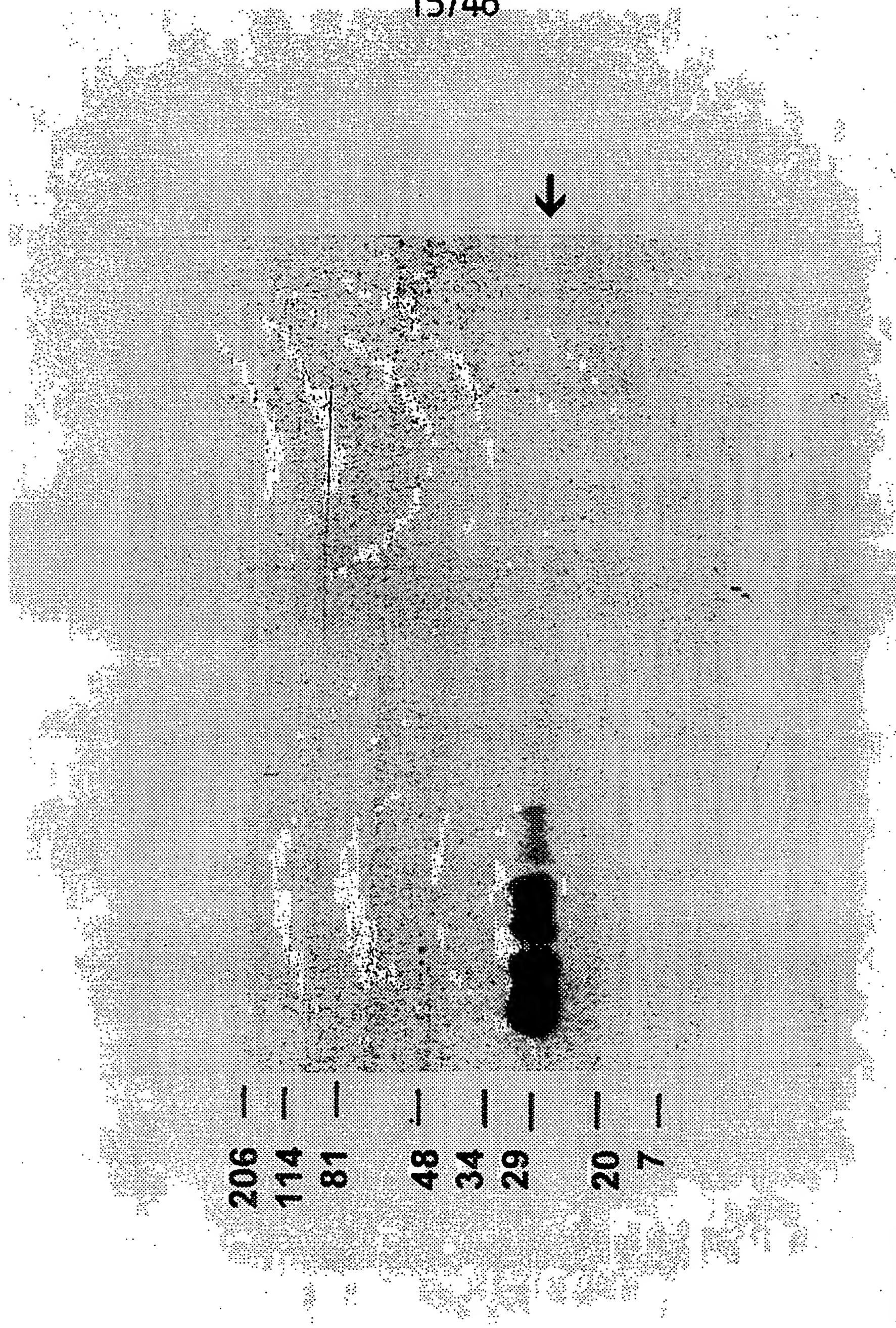
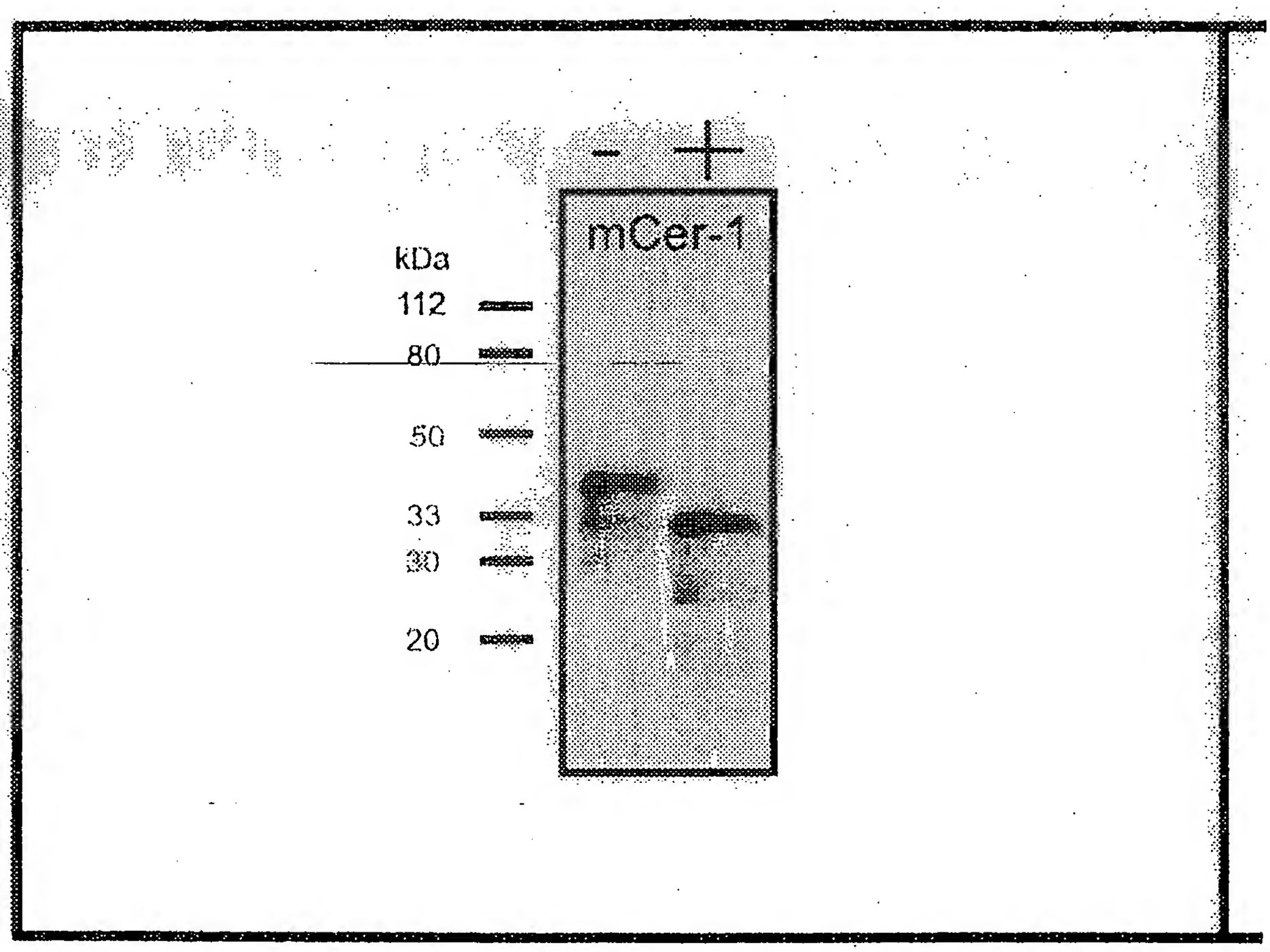


FIG 8

FIG. 9A



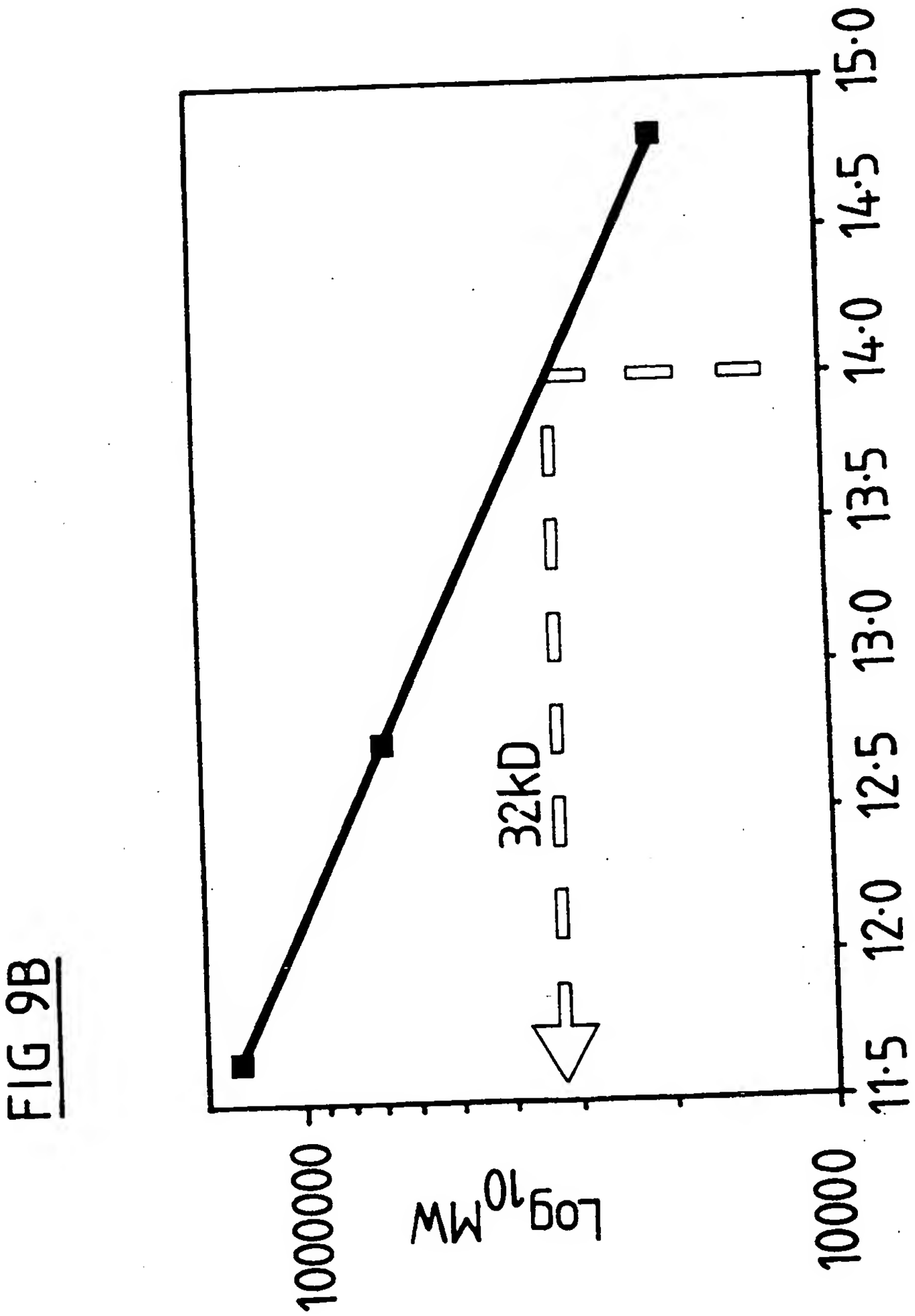

15·0
14·5
14·0
13·5
13·0
12·5
12·0
11·5
Retention time (min.)
32kD
1000000
10000
$\log_{10}$MW

FIG 9B

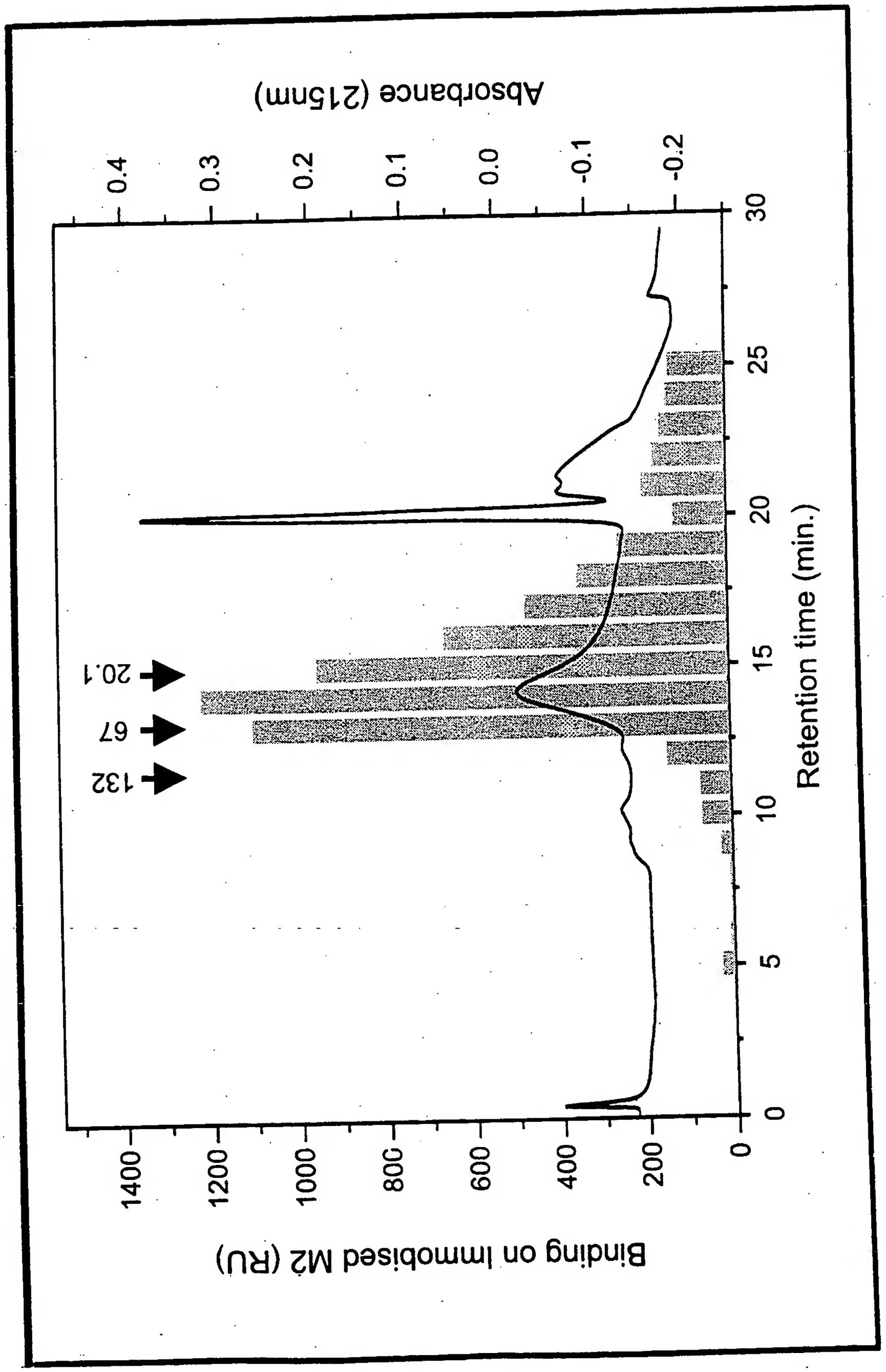

Absorbance (215nm)
0.4
0.3
0.2
0.1
0.0
-0.1
-0.2
30
25
20
15
10
5
0
Retention time (min.)
20.1
67
132
Binding on immobised M2 (RU)
1400
1200
1000
800
600
400
200
0

FIG. 9C

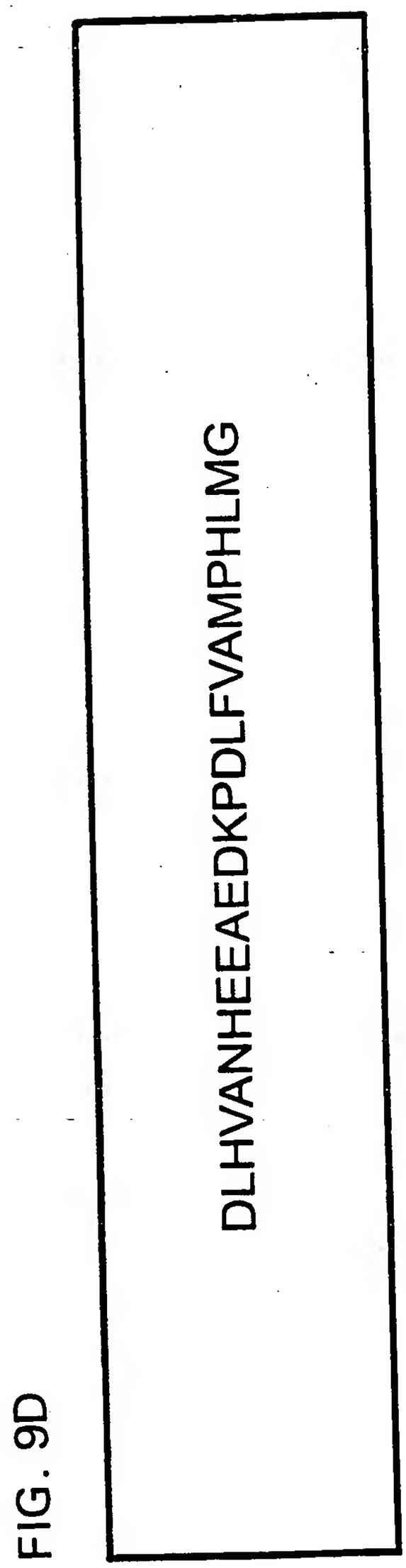


FIG. 9D

FIG. 10A

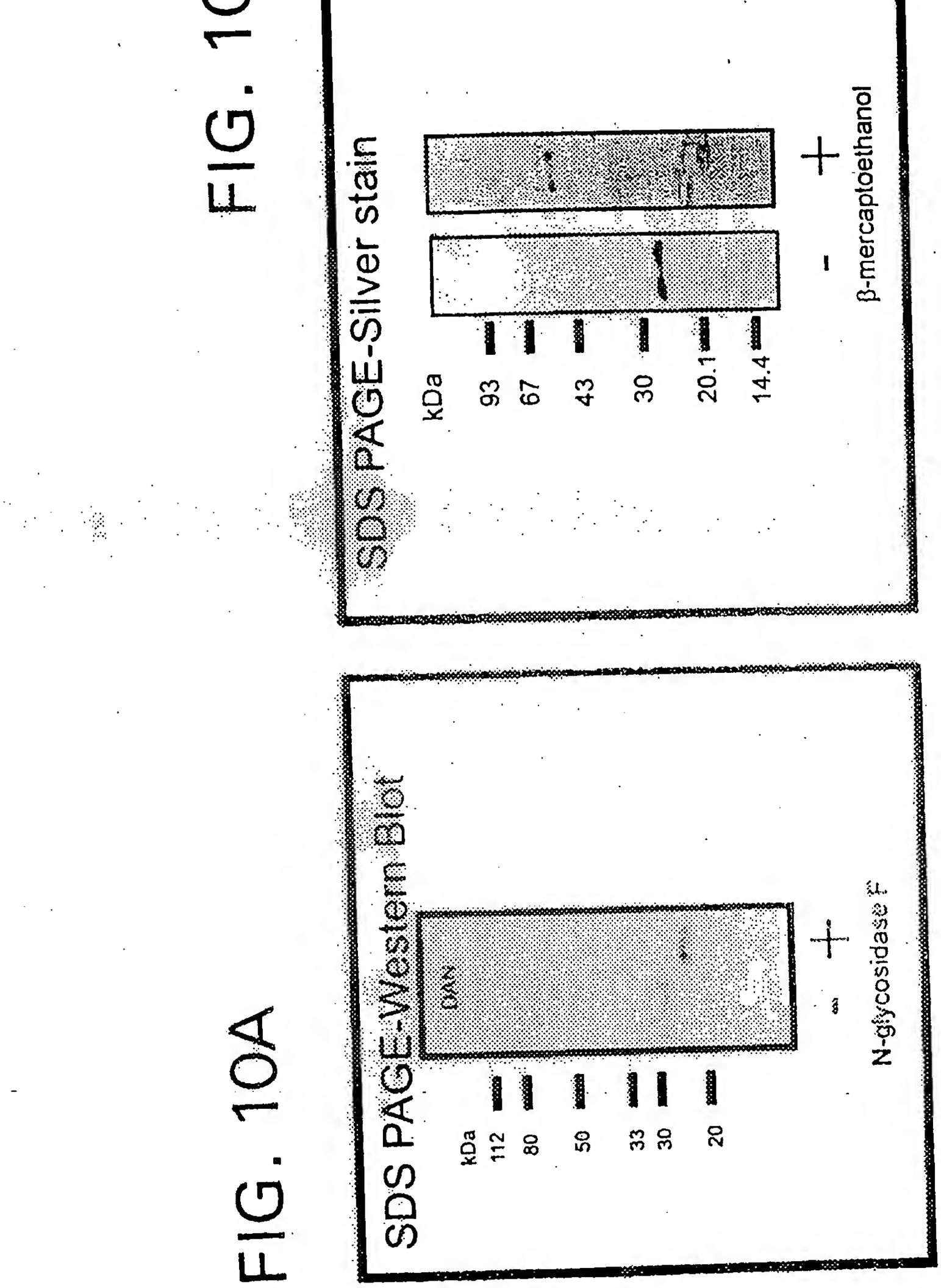


FIG. 10B

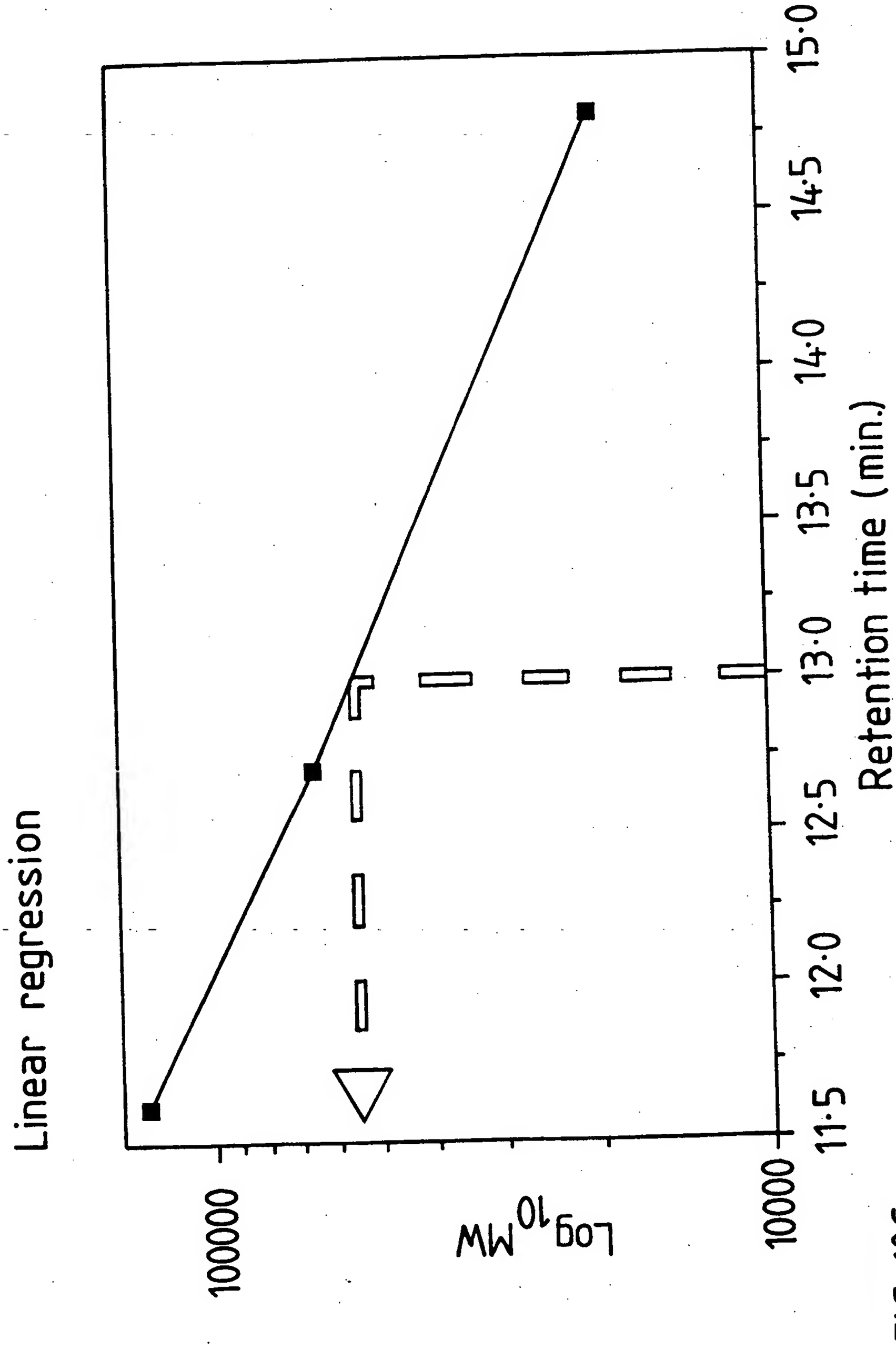

Linear regression
100000
10000
$Log_{10}MW$
12·5
13·0
13·5
14·0
14·5
15·0
Retention time (min.)
11·5
12·0

FIG 10C

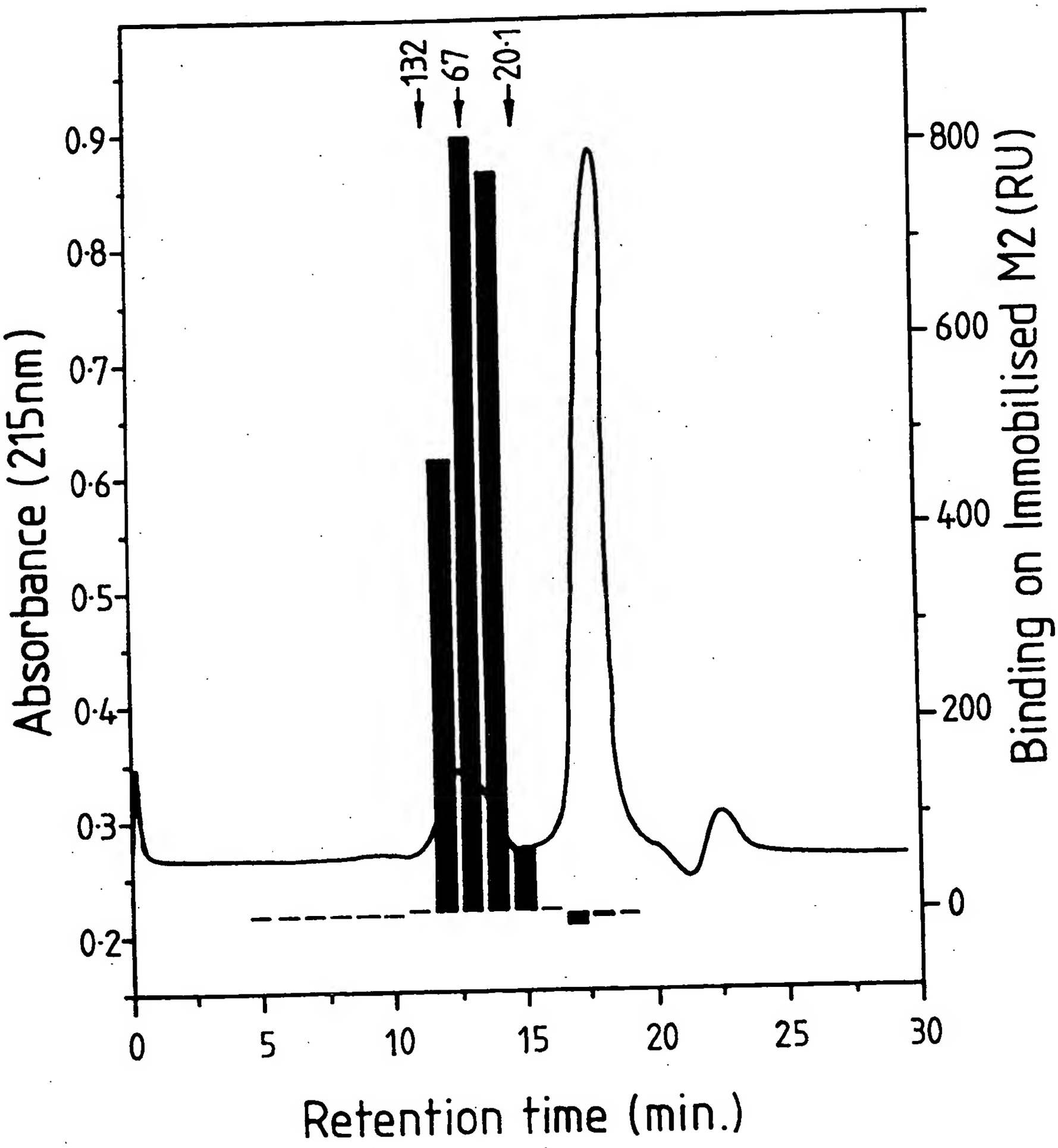

132
67
20·1
Absorbance (215nm)
0·9
0·8
0·7
0·6
0·5
0·4
0·3
0·2
800
600
400
200
0
Binding on Immobilised M2 (RU)
0
5
10
15
20
25
30
Retention time (min.)

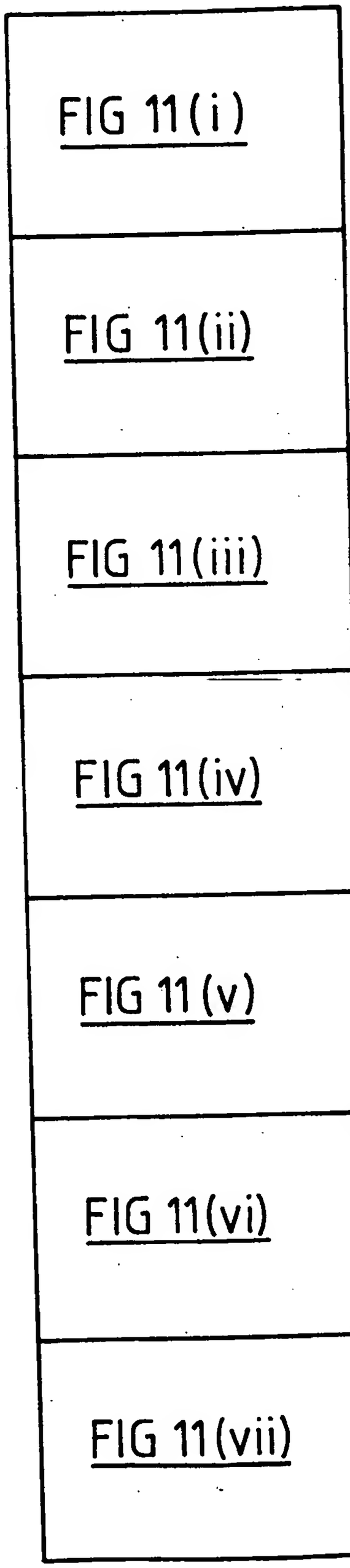
FIG 10D

N terminal sequence analysis: (underlined)

MLWVLVGAVLPVMLLAAPPPINKLALFPDKSAWCEAKNIT 40
QIVGHSGCEAKSIQNRACLGQCFSYSVPNTFPQSTESLVH 80
CDSCMPAQSMWEIVTLECPDHEEVPRVDKLVEKIVHCSCQ 120
ACGKEPSHEGLNVYVQGEDSPGSQPGPHSHAHPHPGGQTP 160
EPEEPPGAPQVEEEGAEDDYKDDDDK 185

FIG 10E

FIG 11



```
TATTCTATTC CAGCACAAGG CAGATGCACA GCAAATGTGA GCTGACTCTA          50

GTCCTTCTTC TGAAAACAGC CATGGGAAAT TTAGGCAAAG AATGTGTTGT         100

CTTTGCTAAT ACTGCTCTTT AAGCCCCAGA CATAGCTAAA CTCTTAGCTA         150

ATTACCCCCT GGGTCCCAGG CTTTCACTGG GGCCTTTTAA AATACACAAA         200

ACCAAAGTGA CGGCAGGAGG CCATTAGCAC TACATAATTC AAGCAAACAA         250

TAAATGTGTT TATTCTGCCT GGCTACTGAC CACCTGCCTT CCCATCCCGC         300

CAGGCAGGTA TCTATATATA CGATTTCCTT TTTCCCAGTC CTGCAGAGAA         350

TGAGCCTCTC CTTTGGGCCT CATCATTTAC AAAAGAAGCT TGGGCCCCTG         400

ACAGC ATG CAT CTC CTC TTA TTT CAG CTG CTG GTA CTC CTG CCT       444
      Met His Leu Leu Leu Phe Gln Leu Leu Val Leu Leu Pro
                          5                  10
                                                  1

CTA GGA AAG ACC ACA CGG CAC CAG GAT GGC CGC CAG AAT CAG AGT    489
Leu Gly Lys Thr Thr Arg His Gln Asp Gly Arg Gln Asn Gln Ser
     15                  20                  25
```

FIG 11(i)

```
TCT CTT TCC CCC GTA CTC CTG CCA AGG AAT CAA AGA GAG CTT CCC   534
Ser Leu Ser Pro Val Leu Leu Pro Arg Asn Gln Arg Glu Leu Pro
     30                  35                  40

ACA GGC AAC CAT GAG GAA GCT GAG GAG AAG CCA GAT CTG TTT GTC   579
Thr Gly Asn His Glu Glu Ala Glu Glu Lys Pro Asp Leu Phe Val
     45                  50                  55

GCA GTG CCA CAC CTT GTA GCC ACC AGC CCT GCA GGG GAA GGC CAG   624
Ala Val Pro His Leu Val Ala Thr Ser Pro Ala Gly Glu Gly Gln
     60                  65                  70

AGG CAG AGA GAG AAG ATG CTG TCC AGA TTT GGC AGG TTC TGG AAG   669
Arg Gln Arg Glu Lys Met Leu Ser Arg Phe Gly Arg Phe Trp Lys
     75                  80                  85

AAG CCT GAG AGA GAA ATG CAT CCA TCC AGG GAC TCA GAT AGT GAG   714
Lys Pro Glu Arg Glu Met His Pro Ser Arg Asp Ser Asp Ser Glu
     90                  95                  100

CCC TTC CCA CCT GGG ACC CAG TCC CTC ATC CAG CCG ATA GAT GGA   759
Pro Phe Pro Pro Gly Thr Gln Ser Leu Ile Gln Pro Ile Asp Gly
     105                 110                 115
```

FIG 11(ii)

```
ATG AAA ATG GAG AAA TCT CCT CTT CGG GAA GAA GCC AAG AAA TTC   804
Met Lys Met Glu Lys Ser Pro Leu Arg Glu Glu Ala Lys Lys Phe
    120                 125                 130

TGG CAC CAC TTC ATG TTC AGA AAA ACT CCG GCT TCT CAG GGG GTC   849
Trp His His Phe Met Phe Arg Lys Thr Pro Ala Ser Gln Gly Val
    135                 140                 145

ATC TTG CCC ATC AAA AGC CAT GAA GTA CAT TGG GAG ACC TGC AGG   894
Ile Leu Pro Ile Lys Ser His Glu Val His Trp Glu Thr Cys Arg
    150                 155                 160

ACA GTG CCC TTC AGC CAG GTATGTGTTC TGGGGGGAGA GCAGGTAAGA      942
Thr Val Pro Phe Ser Gln
    165

GTTTGCAGGT GGTAGTGGAC AGCTGGGATG GATGGAGAGT AGGGGAAAAG        992

GCTGTCAGGA GCCTGACTCT AGCTTAACTA CAGATTTGGT CCTTGGGCAT       1042

TCATCATAGG ATTTGGCAAA GATTAAGTTT CCTTCTGGCC TTTACCATTT       1092

TTTCTTGGCA TTGTGGAAAT GCTGCAAGAA  TGATATGATG ATACTGTCAA      1142
```

FIG 11(iii)

```
TATCAGTAAT CATTCATTCA CACTGAAGAC ACAGAGCTCT GTTTTATTTA      1192

TTTATTTTTG CATTGGAGGT GATCTACTCA GAGATATAAG TCAGACTGTA      1242

CCCTCAGTTA GGAAACTGAG AATTTAGAGT AATCACCAGA ACTCCTCTGT      1292

AGCTATCTTT CTGCACTCTA TTAATATGTG GATGAGCAGG TCAACTCCAT      1342

TTGTTGATAA AGTGGGGTGC ATTGGACTCC TTCCCAAATA CTCTCATATC      1392

CATTTACGAT GGTCTTAATC CCCATAGTCC ATACTTAATT ACTTTATAGG      1442

TTTATGAGGG ACTTCTTTAA TAGCTTGCTA AAGCTATCCC ACAACCTCAA      1492

AGTACGTTGA GGTTCTCAGG CAAAAGTTGT CATATCATTT CTAGTATTAT      1542

GATAGCAAAA AAGTGATTTT CTTTCACTTA TTTTCTCATA TGAGCTTTTT      1592

AAAAAATCAA TCTTGATGTG AGATCATATC TCCTCCCCTT AGAAGTACCT      1642

TTCTCCTGAT TCATGTTGTG TTGGCTGATT TGTAGTTATT ATGATCAATT      1692

CCATGCTATT AAGACAAAGG GACATCCTAC TGTCTACTTC CTCTGGCAAT      1742
```

FIG 11(iv)

```
ATCTACATTC CAAATGTTAA ATTAAAATTG AGAACTTGCA TTAGGTCCTT      1792

AAACATGAAG ATATTGAACC AAAAACATGC AGGGTAGAGT AAAATTTTAT      1842

AGTCGAGTAA TGCTACCCAA TTAAGCAAGC AATAGAATAG GGCAATTGAC      1892

TGTTCAAGGC AGTTAAGTAT TCTGCCTGAA AAGGCAAGGA TATGTAGCAA      1942

TGGCAAGTCA ATTATCAAAT AATAATGACT ACTCTGTTGG CCATGTGCAA      1992

TTAGAAAATT ACCCCTAAGA ATCAGGCAAT CAAATTTCTT TTGAAATTCT      2042

TCTTTTGAAT TCTATTGCTA ATTAAATTAA AACTAAGATG TTTGACTCTT      2092

ACATATTTTG AAAGGCATAT AAAGCTAGGT GCTTGGAGTT ATGAGAGGTA      2142

AAGGTGATGT AATATACAAT GATTTGCAGG CATATGCATT GTAACTCTGC      2192

TTGCATACAA CTTCATAGAC TTGAATGTAC TACAGGTCTT GCAGAATAGG      2242

ATAGAATTAA ACCTAGAATG TTCTGATCTA TTCTACGATC AATGTAACAA      2292

ATATGTATTG GGAGCCTACT ATGCACAAAG CCCTGTGAGG AATAAAAAAG      2342
```

FIG 11(v)

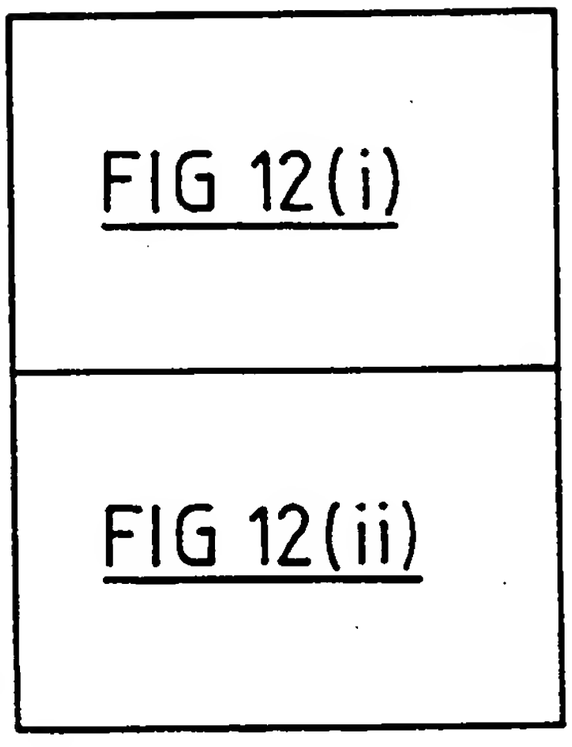
```
TAAGGCACAT TACTTATGTA AGATAATTAC CATTAGAATT TTTCAATCGC        2392

TCACATCCAA TTAGACAAAA TTGCTTAAGG TTTTGCACGA ATAATGTAGA        2442

GTAAAATATT TTTTATGTTA ACTTAGGGAT TCCCTAAAGG CTGTTTAATA        2492

ATTTACTCAA TAAAGAAAAT TTAATTGAGG TGGTTCTGTG CCCTTATAGA        2542

TACCATCACT TGCATATTGC AAATTGTATC CAAAATTGGA AAGCTTTGAA        2592

ATTTTTAAAT TATCCTCAGA TTTACAGTCC ATAGCTTCTG CATTATGTGT        2642

GTTAAAGAAA TAATTCAAAA TAACGTAATG GAAATGTGTT TGCTTTTTAG        2692

ACT ATA ACC CAC GAA GGC TGT GAA AAA GTA GTT GTT CAG AAC AAC    2737
Thr Ile Thr His Glu Gly Cys Glu Lys Val Val Val Gln Asn Asn
  1               5                  10                  15

CTT TGC TTT GGG AAA TGC GGG TCT GTT CAT TTT CCT GGA GCC GCG    2782
Leu Cys Phe Gly Lys Cys Gly Ser Val His Phe Pro Gly Ala Ala
                 20                  25                  30

CAG CAC TCC CAT ACC TCC TGC TCT CAC TGT TTG CCT GCC AAG TTC    2827
Gln His Ser His Thr Ser Cys Ser His Cys Leu Pro Ala Lys Phe
```

FIG 11(vi)

```
 30                  35                  40                  45

ACC ACG ATG CAC TTG CCA CTG AAC TGC ACT GAA CTT TCC TCC GTG 2872
Thr Thr Met His Leu Pro Leu Asn Cys Thr Glu Leu Ser Ser Val
                50                  55                  60

ATC AAG GTG GTG ATG CTG GTG GAG GAG TGC CAG TGC AAG GTG AAG 2917
Ile Lys Val Val Met Leu Val Glu Glu Cys Gln Cys Lys Val Lys
                65                  70                  75

ACG GAG CAT GAA GAT GGA CAC ATC CTA CAT GCT GGC TCC CAG GAT 2962
Thr Glu His Glu Asp Gly His Ile Leu His Ala Gly Ser Gln Asp
                80                  85                  90

TCC TTT ATC CCA GGA GTT TCA GCT TGAAGAGCTA TCCCACTATT       3006
Ser Phe Ile Pro Gly Val Ser Ala

ACCTTTGAAA AGCAAAACCA CAACAGCAAA GATGCTGATT ATTCAGTCTG      3056

AAAATGTTAA GTGGGTACAT AACATTTTCA GGGAAAGGTG ACTTGAAACG      3106

TAGTTTTAAA TTAGAACGAT AGAGGAAATG ATATTAGTCT AGTT            3150
```

FIG 11(vii)

FIG 12



```
CTGCAGAGAA TGAGCCTCTC CTTTGGGCCT CATCATTTAC AAAAGAAGCT          50

TGGGCCCCTG ACAGC ATG CAT CTC CTC TTA TTT CAG CTG CTG GTA        95
                 Met His Leu Leu Leu Phe Gln Leu Leu Val
                   1               5                  10

CTC CTG CCT CTA GGA AAG ACC ACA CGG CAC CAG GAT GGC CGC CAG   140
Leu Leu Pro Leu Gly Lys Thr Thr Arg His Gln Asp Gly Arg Gln
                 15                  20                  25

ACT ATA ACC CAC GAA GGC TGT GAA AAA GTA GTT GTT CAG AAC AAC   185
Thr Ile Thr His Glu Gly Cys Glu Lys Val Val Val Gln Asn Asn
                 30                  35                  40

CTT TGC TTT GGG AAA TGC GGG TCT GTT CAT TTT CCT GGA GCC GCG   230
Leu Cys Phe Gly Lys Cys Gly Ser Val His Phe Pro Gly Ala Ala
                 45                  50                  55

CAG CAC TCC CAT ACC TCC TGC TCT CAC TGT TTG CCT GCC AAG TTC   275
Gln His Ser His Thr Ser Cys Ser His Cys Leu Pro Ala Lys Phe
                 60                  65                  70

ACC ACG ATG CAC TTG CCA CTG AAC TGC ACT GAA CTT TCC TCC GTG   320
Thr Thr Met His Leu Pro Leu Asn Cys Thr Glu Leu Ser Ser Val
```

FIG 12(i)

```
                    75                  80                  85

ATC AAG GTG GTG ATG CTG GTG GAG GAG TGC CAG TGC AAG GTG AAG  365
Ile Lys Val Val Met Leu Val Glu Glu Cys Gln Cys Lys Val Lys
                    90                  95                  100

ACG GAG CAT GAA GAT GGA CAC ATC CTA CAT GCT GGC TCC CAG GAT  410
Thr Glu His Glu Asp Gly His Ile Leu His Ala Gly Ser Gln Asp
                    105                 110                 115

TCC TTT ATC CCA GGA GTT TCA GCT TGAAGAGCTA TCCCACTATT        454
Ser Phe Ile Pro Gly Val Ser Ala
                    120

ACCTTTGAAA AGCAAAACCA CAACAGCAAA GATGCTGATT ATTCAGTCTG       504

AAAA                                                        508
```

FIG 12(ii)

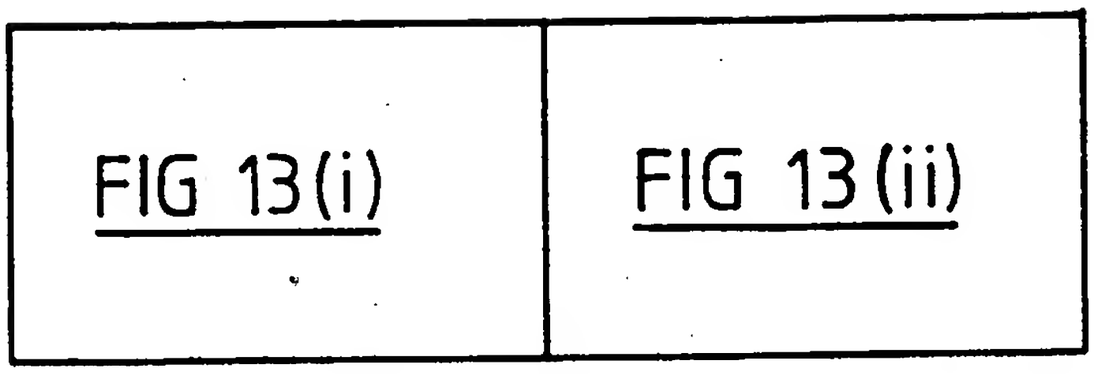


FIG 13

```
Majority        MHLLLXQLLVLLPLGKXXXXXDGXQXQXSLSXXLLXRG--RRE-------
                         10        20        30        40        50
hCRP-1          MHLLLFQLLVLLPLGKTTRHQDGRQNQSSLSPVLLPRN--QRE-------- 41
mouse Cerberus  MHLLLVQLLVLLPLGKADLCVDGCQSQGSLSFPLLERG--RRD-------- 41
X cerberus      MLLNVLRICIIVCLVNDGAGKHSEGRERTKTYSLNSRGYFRKERGARRSK 50

Majority        MLSRXGRFWKKPEXEXXPXRDXXSXXXXSGXQXXXQPXDGXKXEXSPLXE
                        100       110       120       130       140
hCRP-1          MLSRFGRFWKKPEREMHPSRDSDSEPFPPGTQSLIQPIDGMKMEKSPLRE 218
mouse Cerberus  MLSRLGRFWKKPETEFYPPRDVESDHVSSGMQAVTQPADGRKVERSPLQE 218
X cerberus      -MNKVKLFSTVAHGNKSARRKAYNG---SRRNIFSRRSFDKRNTEVTEKP 224

Majority        QTIXHEXCXKVVVQNNLCFGKCXSXHFPGXXQDXHXXCSHCLPXKFTTXH
                        190       200       210       220       230
hCRP-1          QTITHEGCEKVVVQNNLCFGKCGSVHFPGAAQHSHTSCSHCLPAKFTTMH 128
mouse Cerberus  QTIAHEDCQKVVVQNNLCFGKCSSIRFPGEGADAHSFCSHCSPTKFTTVH 128
X cerberus      QNIVHENCDRMVIQNNLCFGKCISLHVP-NQQDRRNTCSHCLPSKFTLNH 135

Majority        XSFIPGXXAXXXXX
                        280
hCRP-1          DSFIPGVSA       267
mouse Cerberus  GSFIPGLPASKTNP 272
X cerberus      --FNMD-TSTTLHH  270
```

FIG 13(i)

```
Majority        ---LXTXNHEEAEXKPDLFVAVPHLXXTSXAGEGQRQRXK
                        60        70        80        90
hCRP-1          ---LPTGNHEEAEEKPDLFVAVPHLVATSPAGEGQRQREK  78
mouse Cerberus  ---LHVANHEEAEDKPDLFVAVPHLMGTSLAGEGQRQRGK  78
X cerberus      ILLVNTKGLDEPHIGHGDFGLVAELFDSTRTHTNRKEPD-  89

Majority        EAKXFWHXFMFRKXPAXQGVILPIKSHEVHWETCRTVPFX
                       150       160       170       180
hCRP-1          EAKKFWHHFMFRKTPASQGVILPIKSHEVHWETCRTVPFS 168
mouse Cerberus  EAKRFWHRFMFRKGPAFQGVILPIKSHEVHWETCRTVPFN 168
X cerberus      GAKMFWNNFLVKMNGAPQNTSHGSKAQEIMKEACKTLPFT 175

Majority        LXLNCTXXXXVVKVVMXVEECQCXVKTEXXXXXXLHAGSQ
                       240       250       260       270
hCRP-1          LPLNCTELSSVIKVVMLVEECQCKVKTEHEDGHILHAGSQ 258
mouse Cerberus  LMLNCTSPTPVVKMVMQVEECQCMVKTERGEERLLLAGSQ 258
X cerberus      LTLNCTGSKNVVKVVMMVEECTCEAHKSN-----FHQTAQ 259
```

FIG 13(ii)

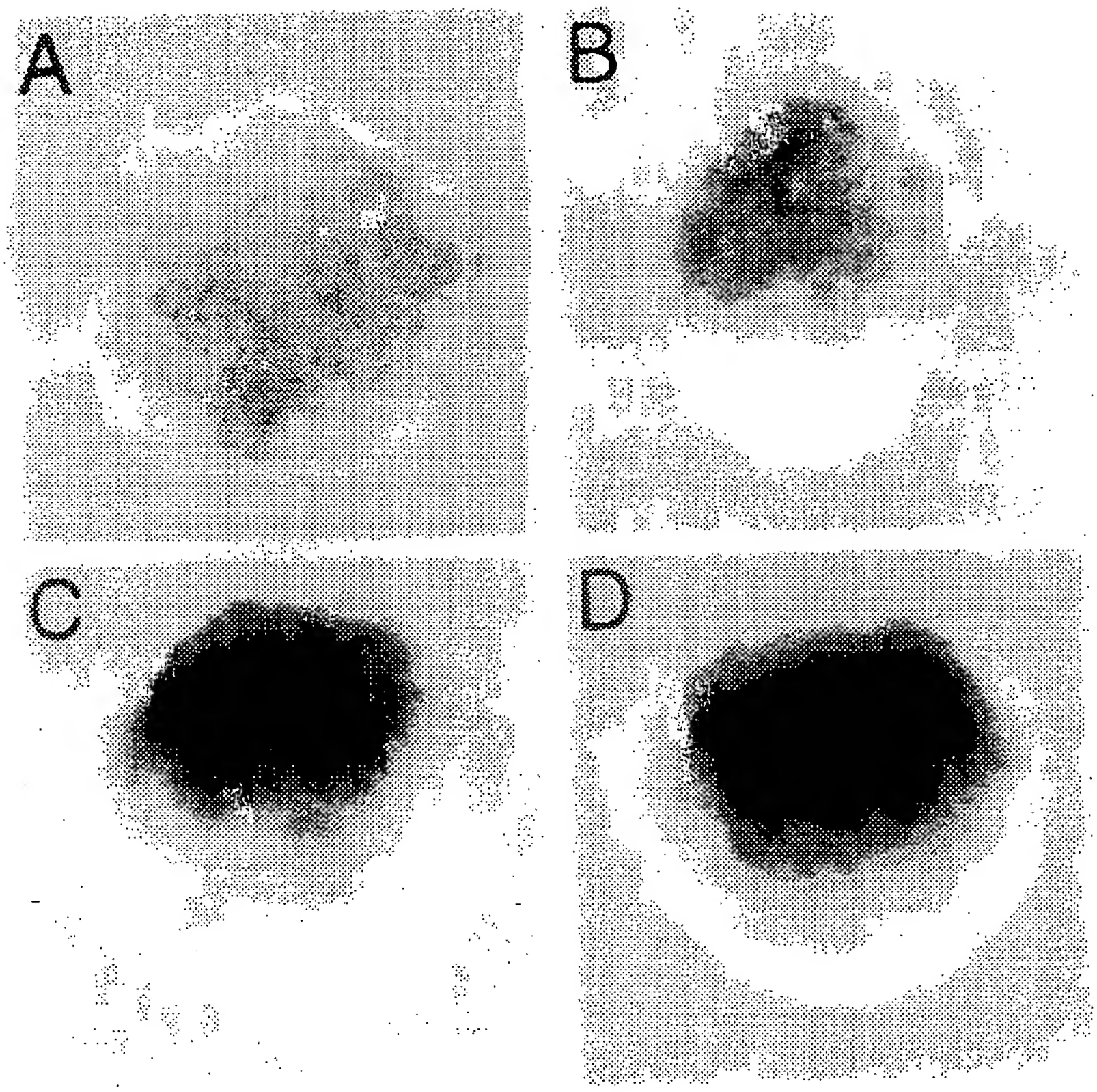
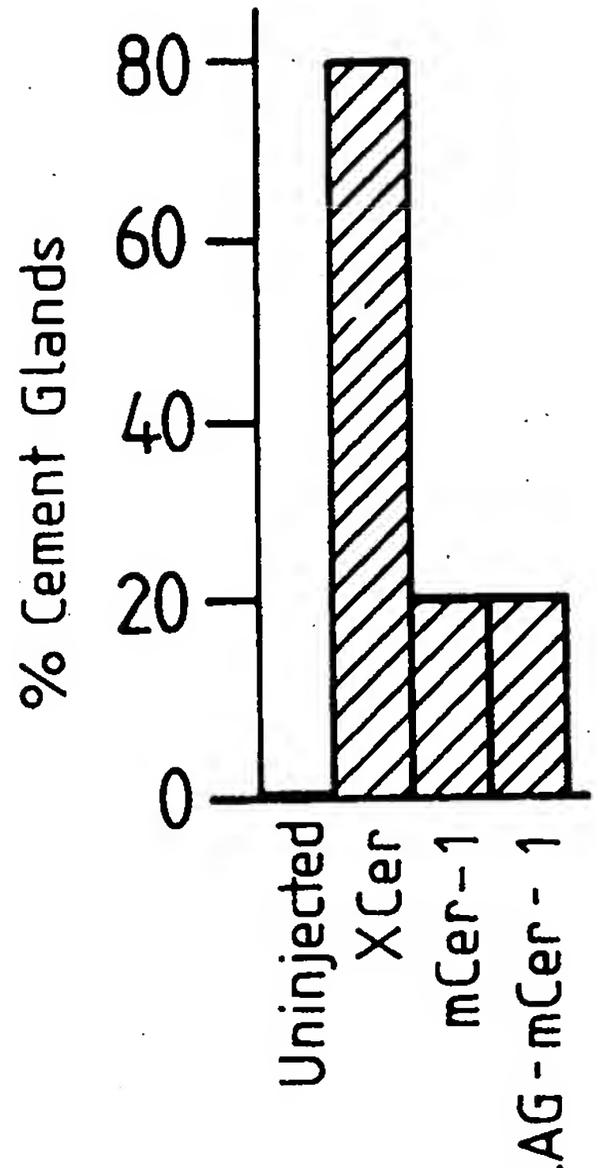


FIG 14A

FIG 14B


% Cement Glands
0
20
40
60
80
Uninjected
XCer
mCer-1
AG-mCer-1
CFL

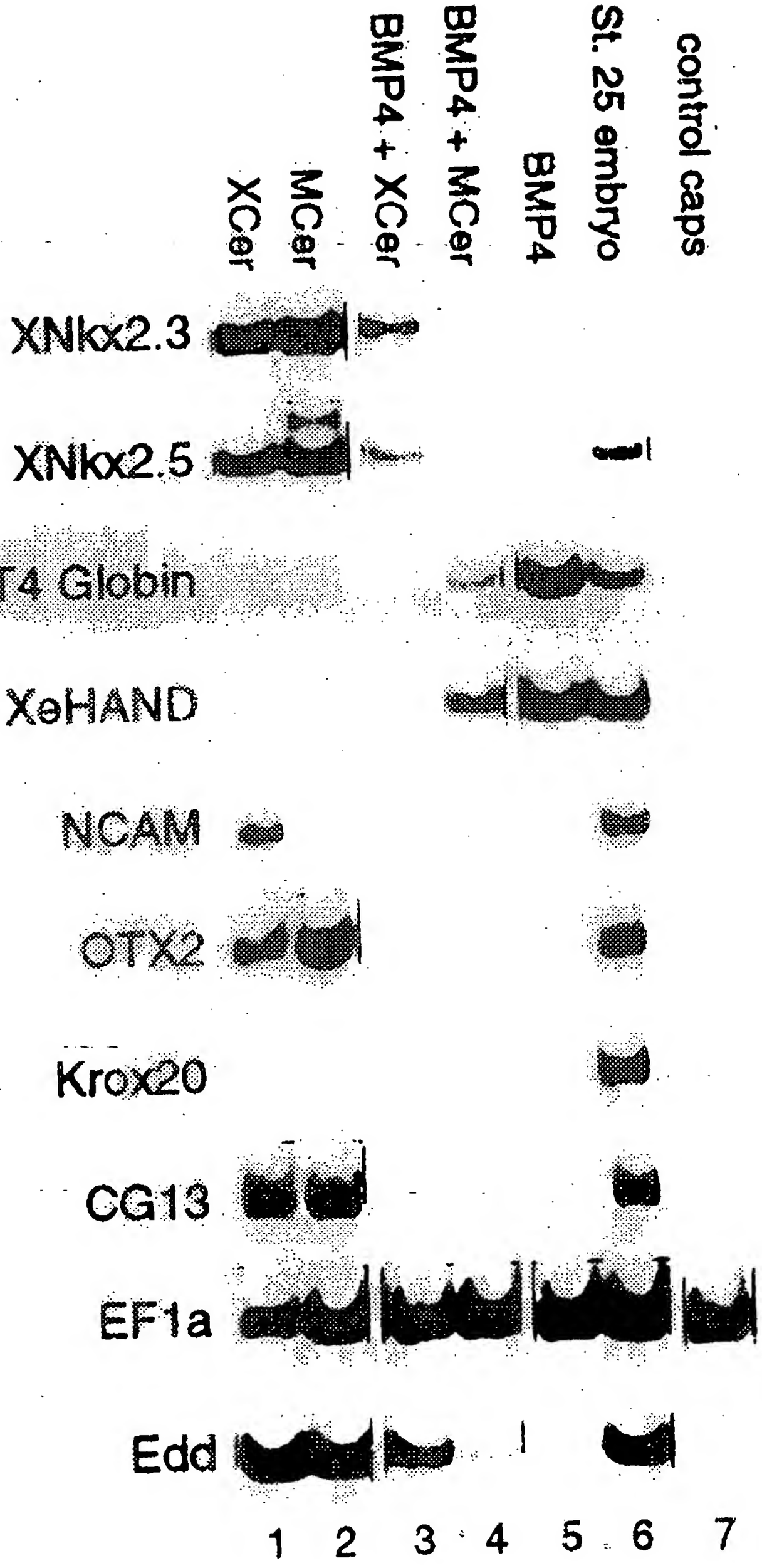

control caps
St. 25 embryo
BMP4
BMP4 + MCer
BMP4 + XCer
MCer
XCer
XNkx2.3
XNkx2.5
T4 Globin
XeHAND
NCAM
OTX2
Krox20
CG13
EF1a
Edd
1 2 3 4 5 6 7

FIG 14C

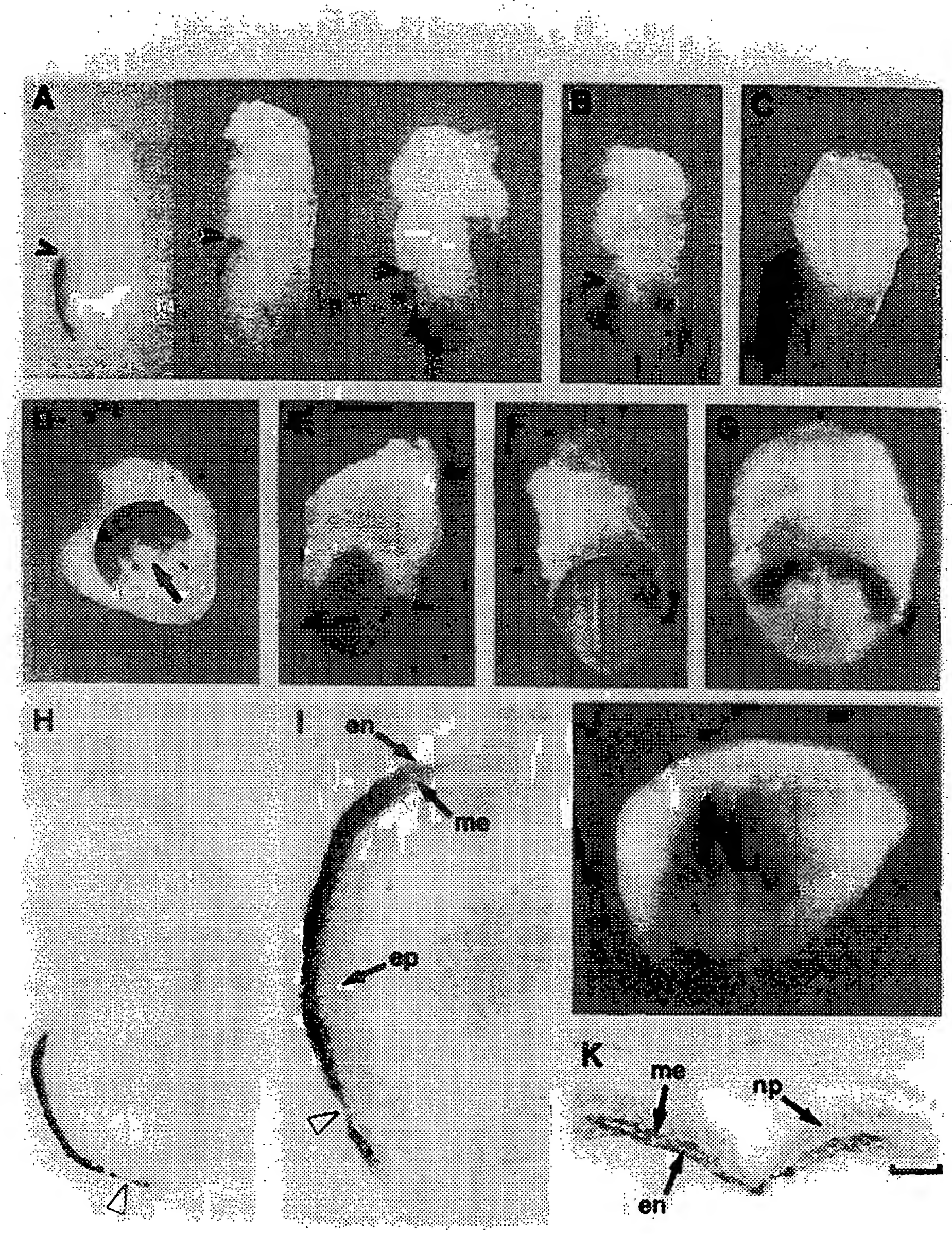


FIG 15

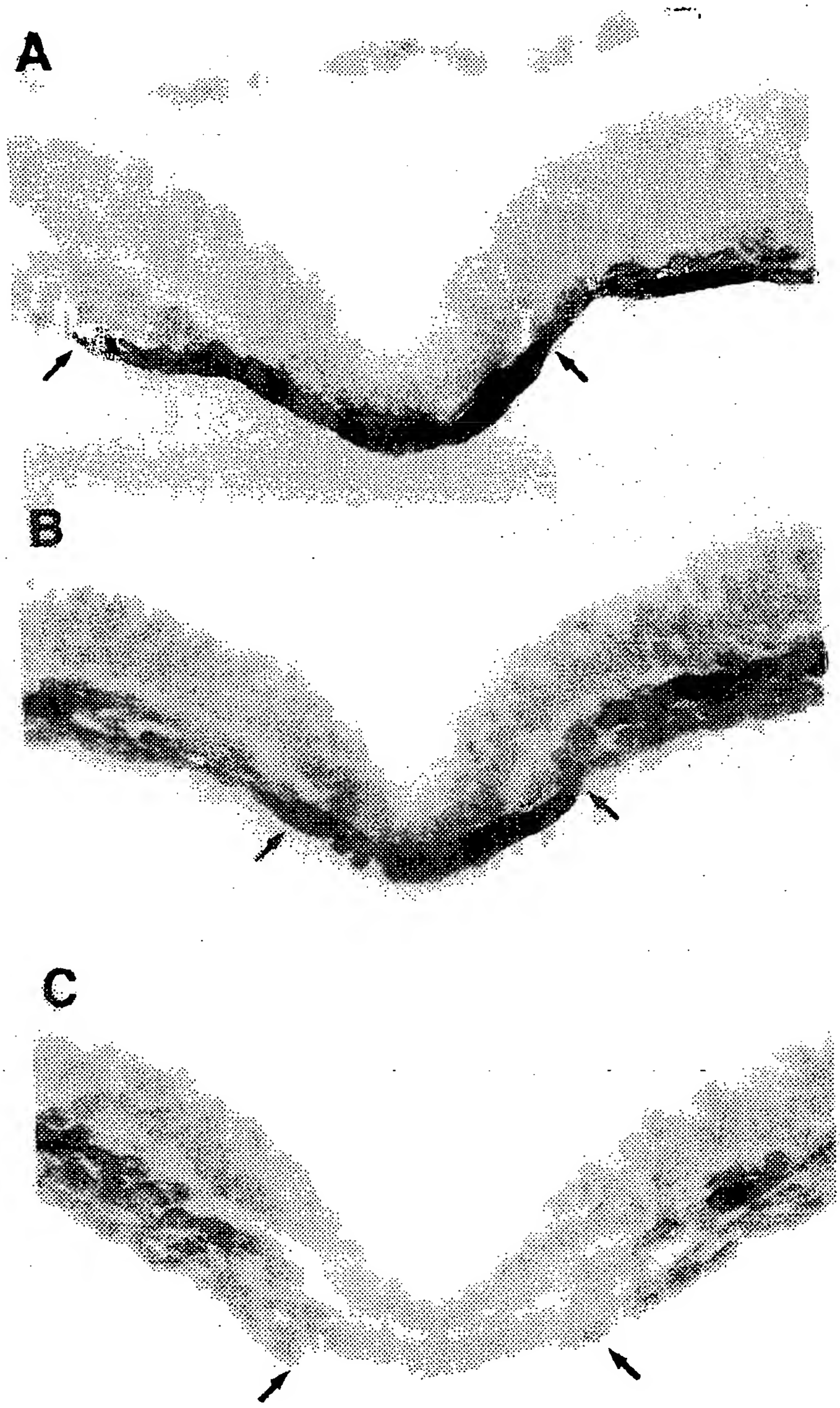


FIG 16

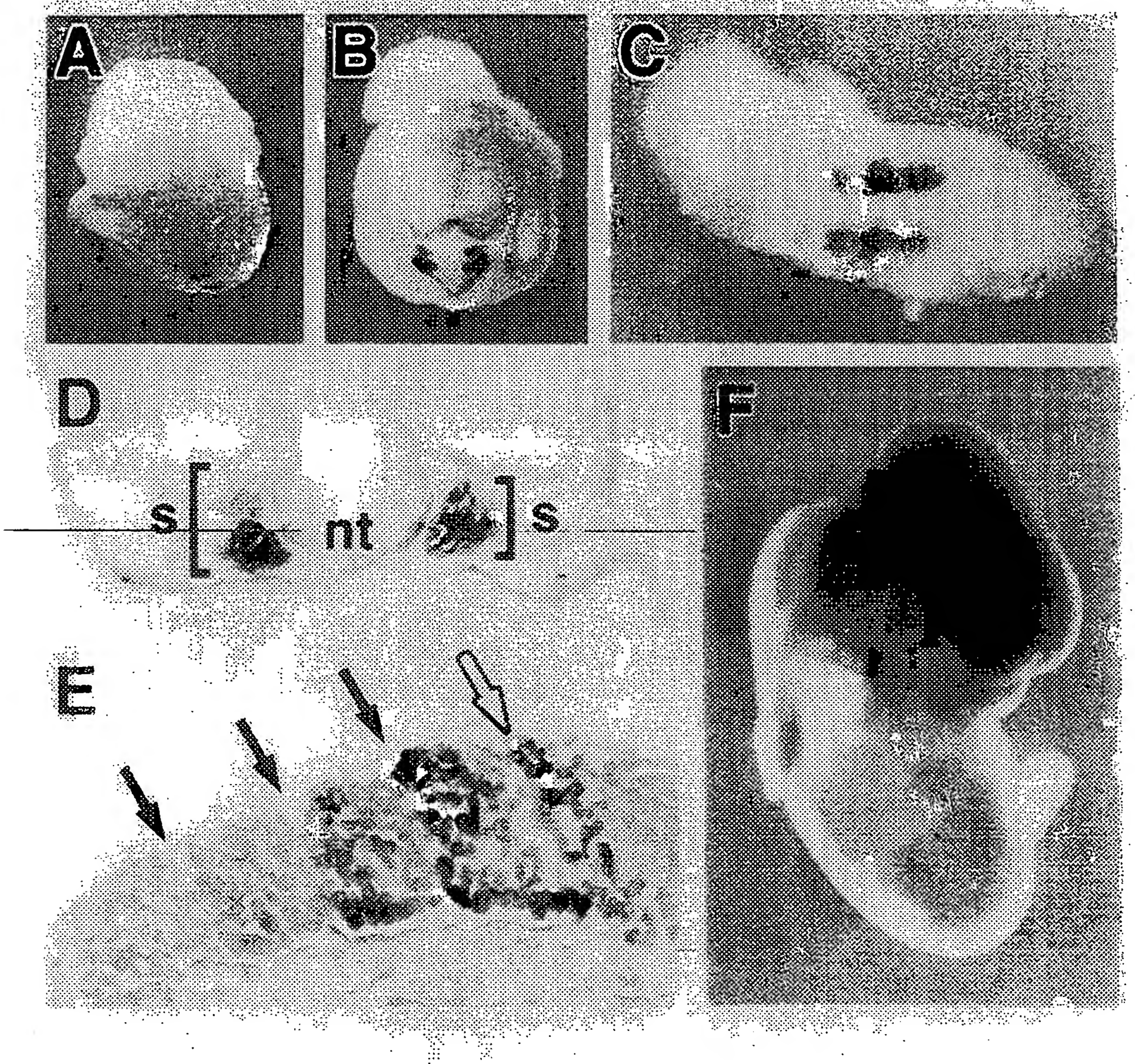


FIG 17

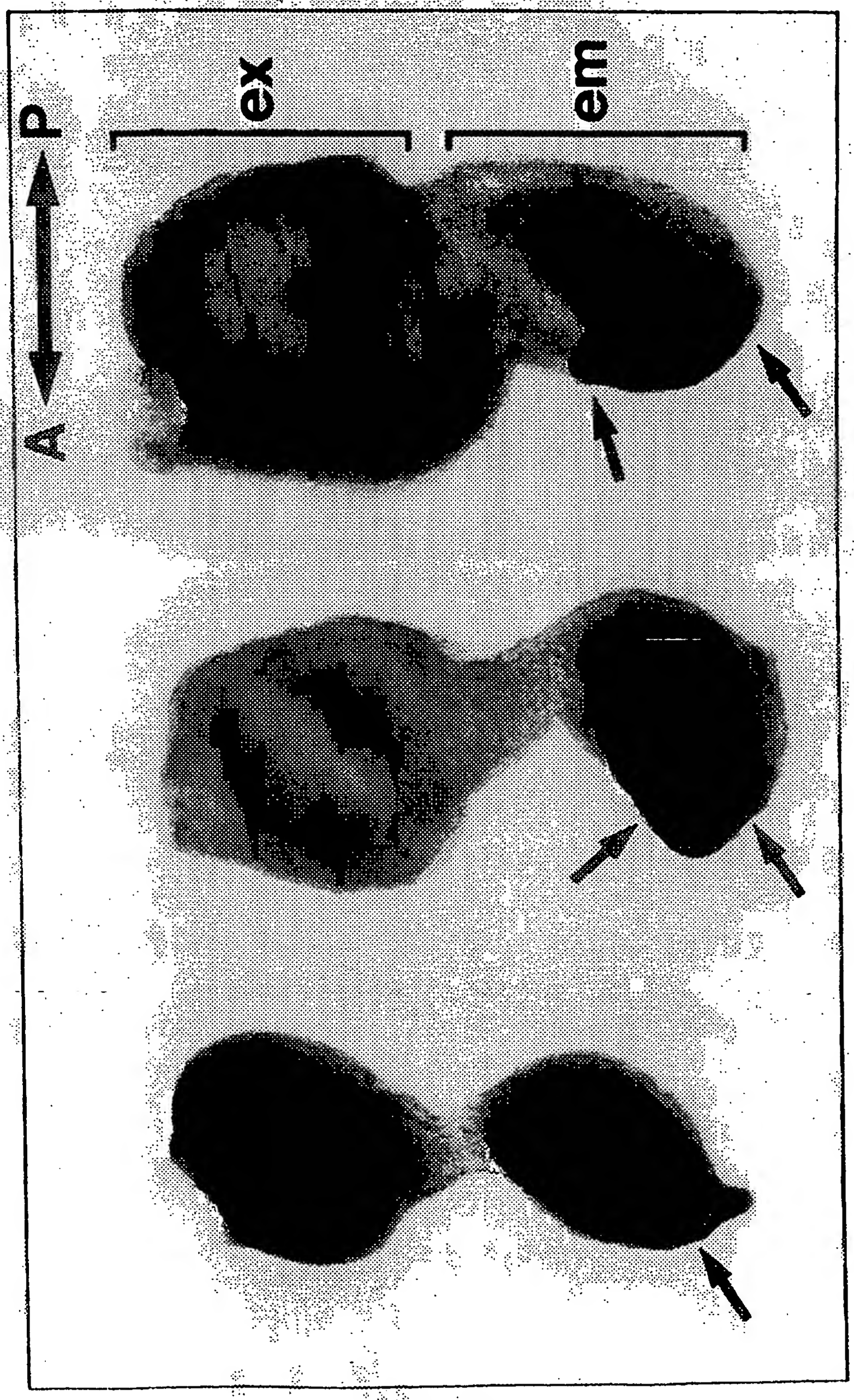


FIG 18

FIG 19A

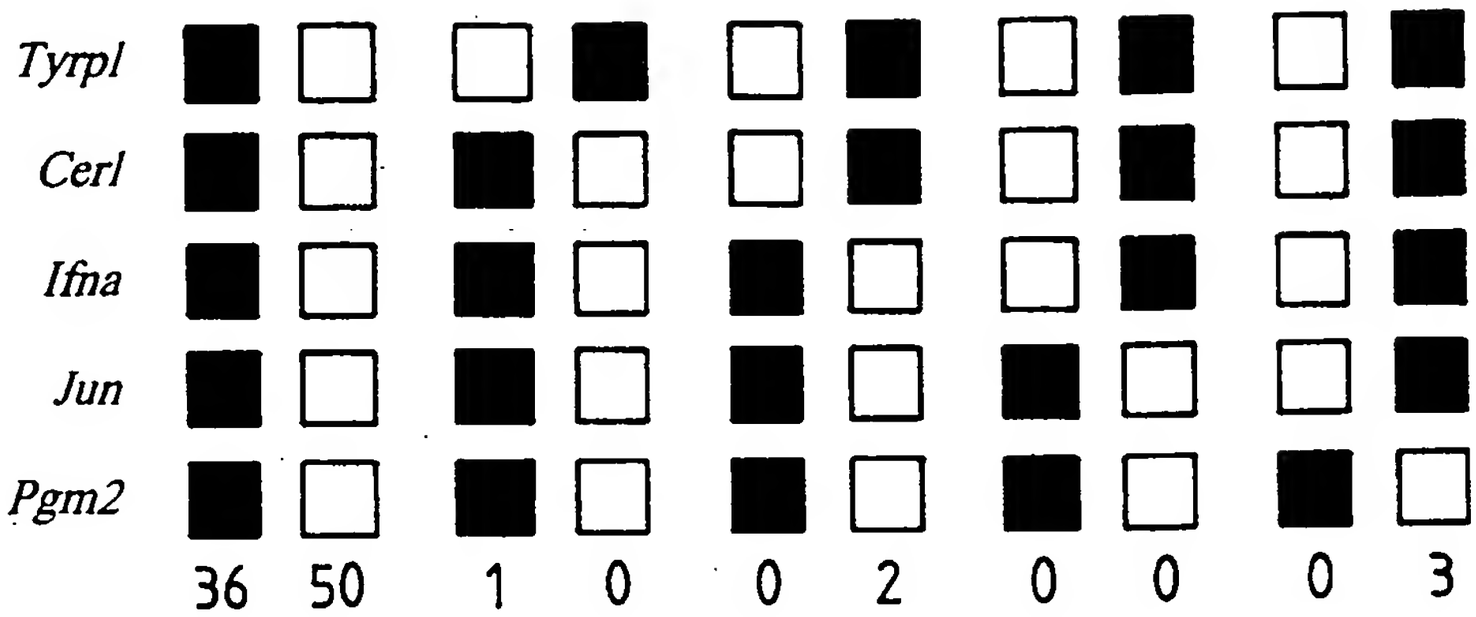
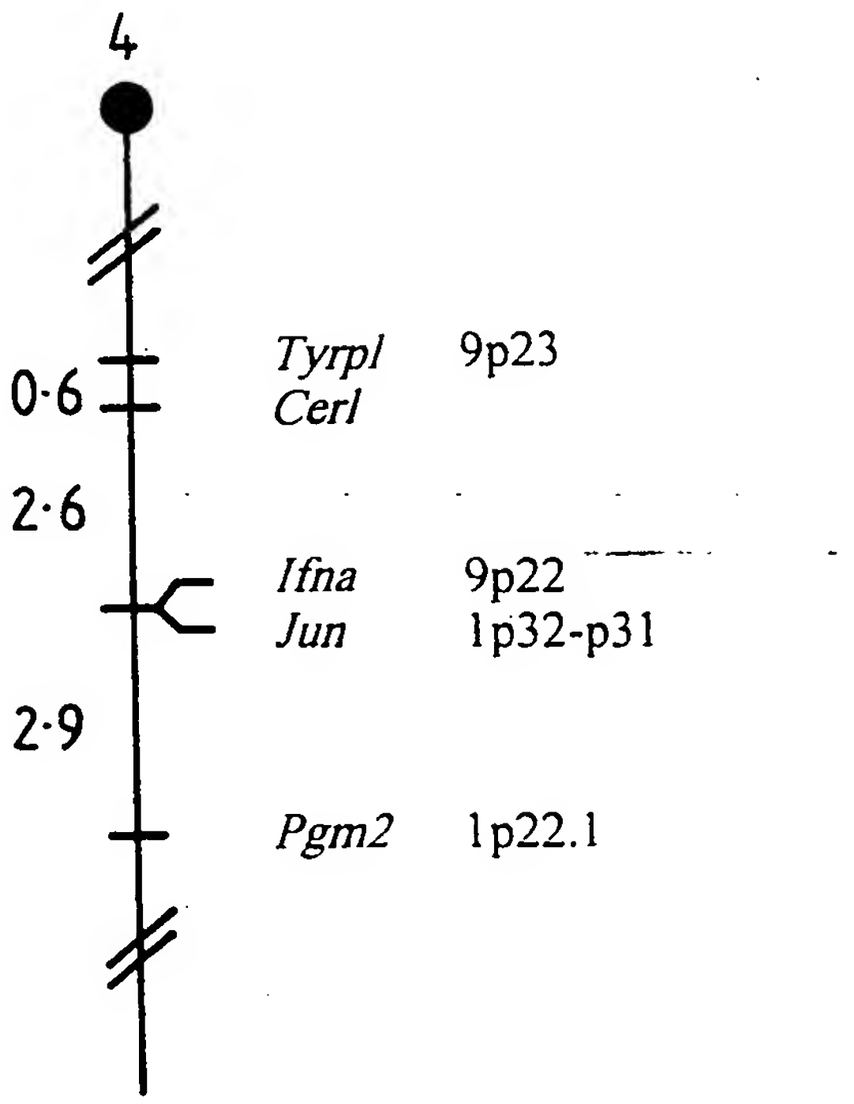


FIG 19B



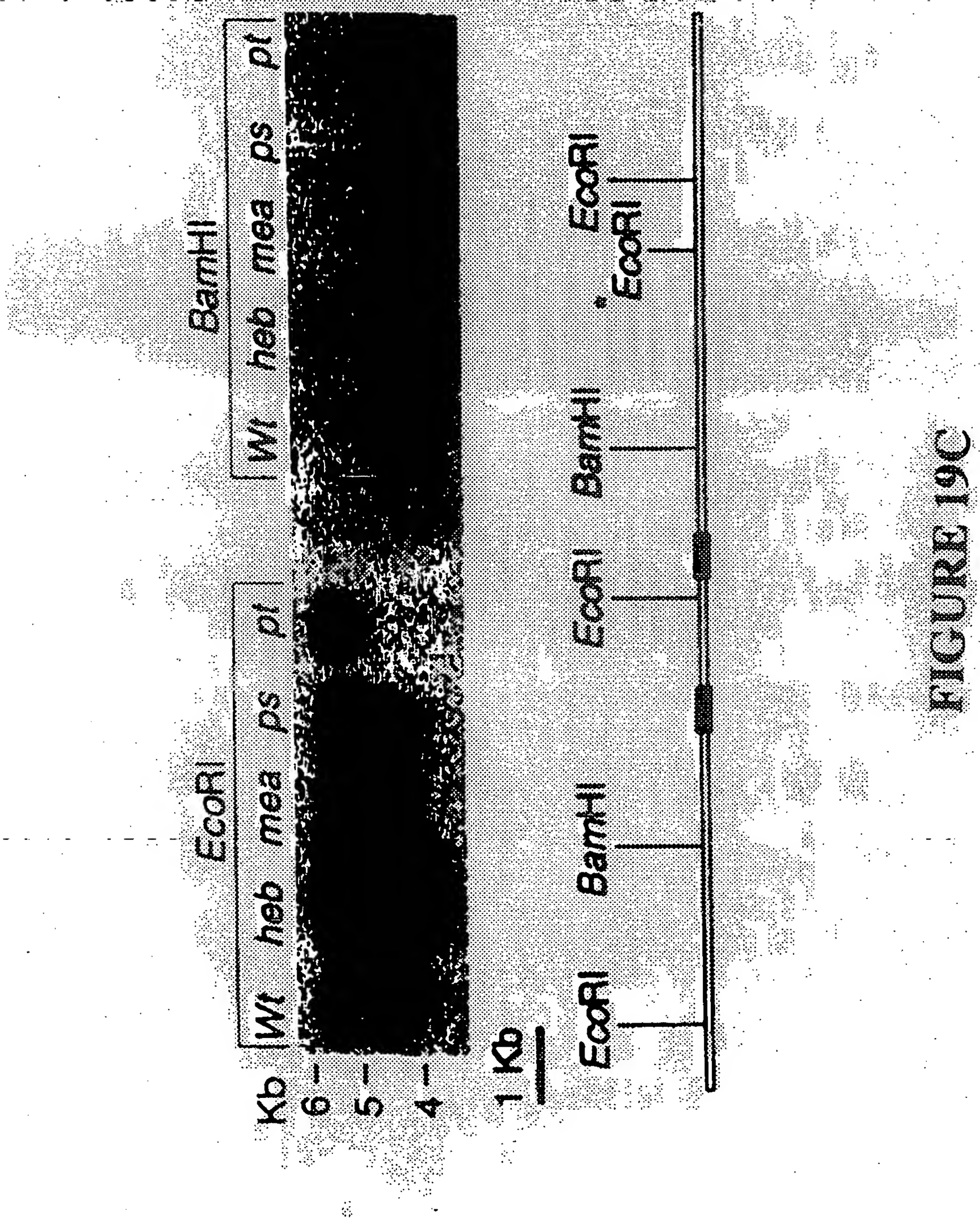

EcoRI
Wt
hsb
mea
ps
pt
BamHI
EcoRI
BamHI
EcoRI
EcoRI
BamHI
Kb
6
5
4
1 Kb

FIGURE 19C

FIG 20

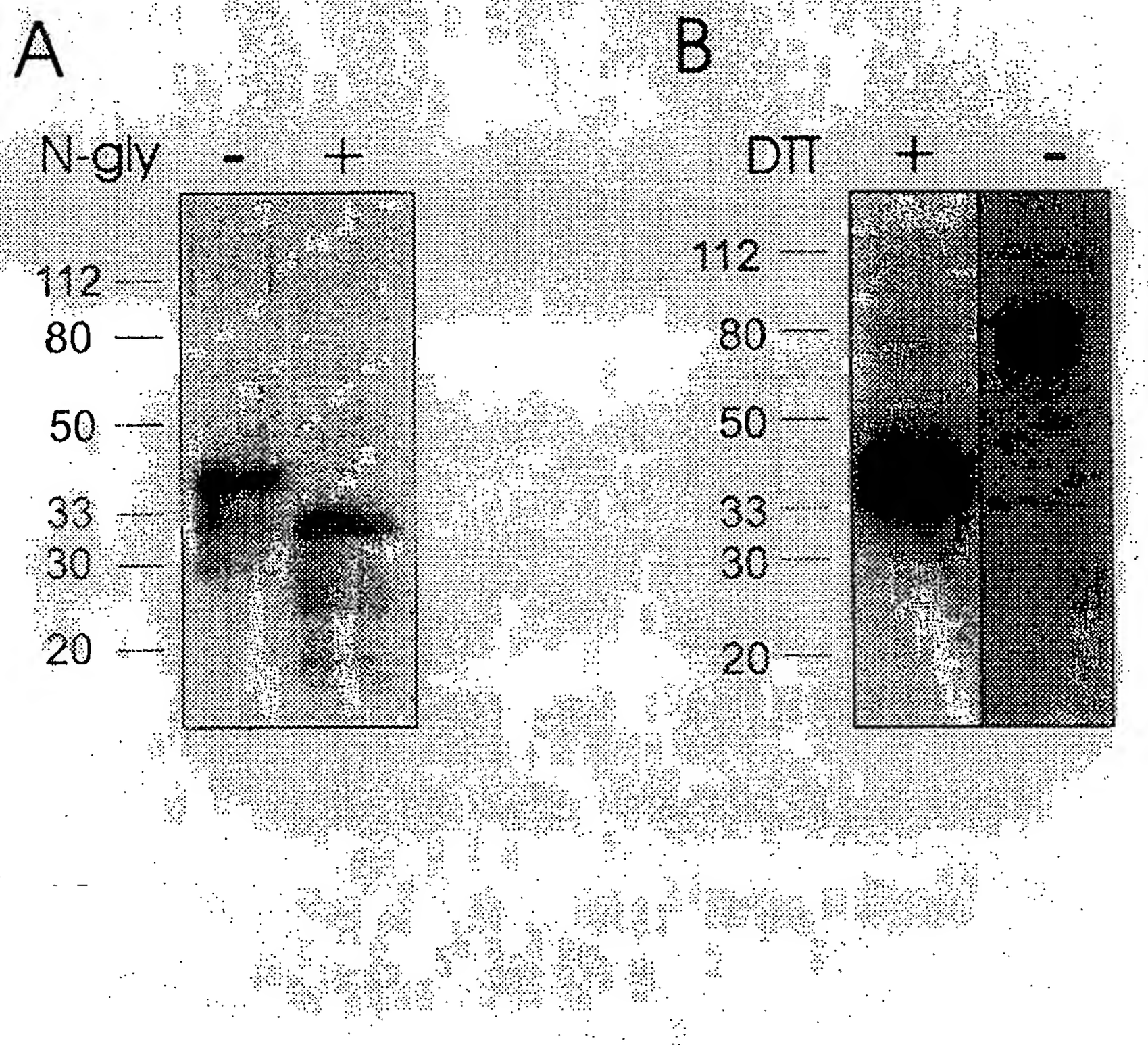
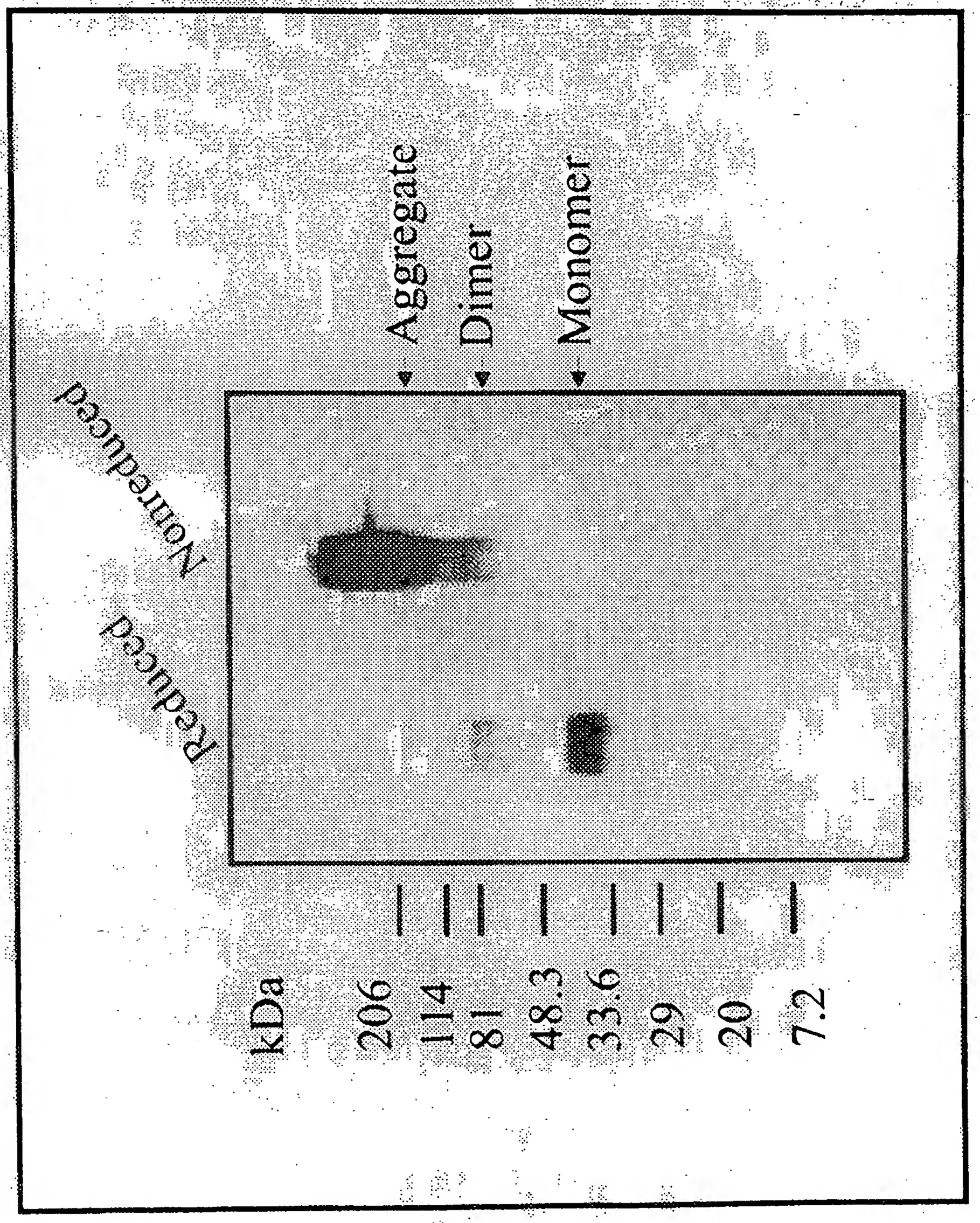


FIG. 21


Aggregate
Dimer
Monomer
Reduced
Non-reduced
kDa
206
114
81
48.3
33.6
29
20
7.2

INTERNATIONAL SEARCH REPORT

International Application No.
PCT/AU 98/00078

A. CLASSIFICATION OF SUBJECT MATTER

Int Cl[6]: C07K 14/52, A61K 38/19

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
As below

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
see attached sheet

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P	WO 97/48275 (THE REGENTS OF THE UNIVERSITY OF CALIFORNIA) 24 December 1997 see Figure 1	1-5,7-9,12-16,18-20,23-27
X	Chemical Abstracts 125:217338 & Nature volume 382 No: 6592 issued (1996) Bouwmeester, T et al "Cerberus is a head-inducing secreted factor expressed in the anterior endoderm of Spemann's organizer page 595-601 see abstract	1-27

[X] Further documents are listed in the continuation of Box C

[X] See patent family annex

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
17 June 1998	22 JUN 1998
Name and mailing address of the ISA/AU AUSTRALIAN PATENT OFFICE PO BOX 200 WODEN ACT 2606 AUSTRALIA Facsimile No.: (02) 6285 3929	Authorized officer **K.F. PECK** Telephone No.: (02) 6283 2263

INTERNATIONAL SEARCH REPORT

International Application No.
PCT/AU 98/00078

C (Continuation) **DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Chemical Abstract 124:142381 & Japan J Cancer Research Volume 87 No: 1 issued (1996) Ozaki, T et al "Cloning of mouse DAN cDNA and its down-regulation in transformed cells" pages 58-61 see abstract	1-27
X	Chemical Abstract 122:179709 & Adv Enzyme Regul Volume 34 issued (1994) Sakiyama, S et al "Molecular cloning and characterization of a cDNA showing tumor-suppressive activity in v-src-transformed 3Y1 rat fibroblasts" pages 247-255 see abstract	1-27
X	Chemical Abstract 122:2407 & Oncogene Volume 9 No: 10 issued (1994) Enomoto, H et al "Identification of human DAN gene mapping to the putative neuroblastoma tumor suppressor locus" pages 2785-2791 see abstract	1-27
X	Chemical Abstracts 119:21591 & Proc Natl Acad Sci USA Volume 90 No: 7 issued (1993) Ozaki, T et al "Molecular cloning and characterization of a cDNA showing negative regulation in v-src-transformed 3Y1 rat fibroblasts" pages 2593-2597 see abstract	1-27
Y	Journal of Biological Chemistry Volume 272 No: 5 issued (31 January 1997) Nakao, A et al "Identification of Smad 2, a human mad-related protein in the transforming growth factor beta signalling pathway" pages 2896-2900 see discussion pages 2899-2900	1,5,12,16,23-27
Y	Current opinion in Genetics and Development Volume 6 No: 4 issued (August 1996) Hogan, B L "Bone morphogenetic proteins in development" pages 432-8 see abstract	1,5,12,16,23-27
A	Medline 97180932 & EMBO Journal Volume 16 No: 2 issued (15 January, 1997), Guenda, A et al "Activation of stress-activated protein kinase-3 (SAPK3) by cytokines and cellular stresses is mediated via SAPKK3 (mKK6); comparison of the specificities of SAPK3 and SAPK2" pages 295-305 see abstract	1-27
A	Medline 95408268 & Biochemical and Biophysical Research Communications Volume 214 No: 2 issued (14 September 1995) Tsujimura, A et al "Developmental and differential regulations in gene expression of Xenopus pleiotrophic factors - alpha and - beta" pages 432-439 see abstract	1-27
A	Roux's Arch Dev Biol Volume 205 No: 5-6 issued (1996) Nakamura, H et al "Isolation of Xenopus HGF gene promoter and its functional analysis in embryos and animal caps" pages 300-310 see discussion on pages 308-309	1,5,12,16,23-27

INTERNATIONAL SEARCH REPORT

International Application No.
PCT/AU 98/00078

Box 1 Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically: please see comments for explanation in the continuation of Box 1.

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a)

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

Supplemental Box
(To be used when the space in any of Boxes I to VIII is not sufficient)

Continuation of Box No: Electronic Data Base

Electronic data base consulted.

Medline: cerberus
CA : xenopus and cytokine
Orbit : signal(s) sequence and cyst:ne

STN Peptide sub sequence search
STN QNNLCFGKC/SIQNRACLGQC/QGVILPILSHEVHWETCRTVPF

INTERNATIONAL SEARCH REPORT

International Application No.
PCT/AU 98/00078

Box I (continued)

1 Claims 1, 11 and 23-27

No meaningful search could be conducted as the terms "signal sequence", "cystine knot domain" and "cerberus related protein" or "CRP" are not effective keywords (search terms).

2 Claims 2 and 13

No meaningful search could be conducted due to the large number of variables, as indicated by {AA} (any amino acid from 0 to 50 residues).

3 Claims 4, 5-11, 14 and 16-22

No meaningful search could be conducted where only partial similarity is needed to the given amino acid sequence.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International Application No.
PCT/AU 98/00078

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document Cited in Search Report		Patent Family Member	
WO,A	97/48275	AU	35765/97

END OF ANNEX